On the time-varying links between oil and gold: New insights from the rolling and recursive rolling approaches

Mehmet Balcilar[1,2,3], Zeynel Abidin Ozdemir[4] and Muhammad Shahbaz[2]

[1] Department of Economics, Eastern Mediterranean University, Famagusta, Turkey

[2] Energy and Sustainable Development, Montpelier Business School, Montpelier, France

[3] Department of Economics, University of Pretoria, Pretoria, South Africa

[4] Department of Economics, Gazi University, Ankara, Turkey

*Correspondence
Zeynel Abidin Ozdemir, Gazi University, Ankara, Turkey. Email: zabidin@gazi.edu.tr

Abstract

This study analyses the dynamic linkages between oil and gold prices for the spot and 1- to 12-month futures markets using monthly data over the period 1983–2016. To do this, we use the rolling and recursive rolling Granger causality approaches. The distinguishing feature of this study from the previous studies is that this is the first study investigating the causal links between oil and gold using time-varying causality tests. The findings show that the causality links between oil and gold display strong time variation. Although causal links are not detected for most of the study period, strong bidirectional or unidirectional causality is found in several subsamples. The duration of the periods with causality links varies from a few months to 3 years, whereas the duration for the noncausality periods might be 15 years long. By date stamping the causality links between oil and gold, our paper discovers that causality from oil to gold is related to large oil price changes, whereas causality from gold to oil is related to large financial crises. The evidence obtained in the paper points out the dangers of assuming a constant causality link between oil and gold markets because these links might break down unexpectedly. Our findings point out to the dangers of assuming noncausality between oil and gold particularly in hedging oil price risk using gold.

Keywords: gold and oil prices, recursive rolling estimation, rolling estimation, time-varying Granger causality

1. Introduction

Oil and gold, which are frequently tradable and have high liquidity and synchronization in their movements, are the two commodities with an irreplaceable role in an economy. Oil is the most traded commodity in the world and has a highly volatile price. Gold, on the other hand, with the lowest volatility, is the leader of all precious metals. When stock market risks are considered, gold is known as a safe haven and used as an efficient hedge instrument (Junttila, Pesonen, & Raatikainen, 2018). As oil has been increasingly becoming financialized, whether gold does also function as an efficient hedge instrument against oil price risk is interest to financial industry. A long-term analysis of the behavioural tendency of oil and gold demonstrates that both commodities historically tend to move simultaneously upward and downward in price, and a change in the price of one appears to induce a change in the price of the other. An observation of the two commodities for the last 50 years, for instance, has shown that they tend to behave in similar ways with respect to price, with a positive correlation of 80% (see Tiwari & Sahadudheen, 2015)—a view that is also supported by a large number of studies. Accordingly, Sari, Hammoudeh, and Soytas (2010) contend that changes in the price of gold mainly stem from fluctuations in the price of oil. Hence, this study empirically investigates the time-varying Granger causality between the oil and gold markets.

The relationship between the oil price and gold price can arise from several channels. Most oil-importing countries' preference to pay for their oil supplies in gold and the investment of a large part of oil-exporting countries' revenues in gold are just a few of the examples to support this suggestion. Still another example is that an increase in oil price means an increase in the cost of gold extraction, and this results in a reduction in the profit margin. It may be said that oil prices are inversely proportional to the share prices of gold mining companies.

Another channel—indeed, the best one according to Narayan, Narayan, and Zheng (2010)—used to explain the link between oil and gold stems from inflation. Accordingly, an increase in the international crude oil price leads to an increase in general price levels due to an increase in transportation and production costs. This causes negative effects on oil importing countries. Therefore, there is a positive relationship between oil price and inflation. Because gold is a unique instrument of inflation hedging in the long term, investors usually invest in gold during periods of increased inflation to balance their portfolio (Ghosh, 2011). The gold price will rise in these periods of high inflation; thus, positive relationships will be observed between oil and gold. The World Gold Council holds the view that gold has always held itself against inflation throughout history. A higher oil price, on the other hand, is likely to fit the bill of countries—net oil-importing countries in particular—and this causes an increase in import costs, which causes a high trade deficit; this in turn influences the value of the domestic currency and the money in circulation, and all these factors lead to inflation—a chain reaction. In this case, gold will increase in response to inflation, which is caused by rising oil prices. Gold, on the other hand, might increase when the goods and financial markets are strong, proxying the market conditions. Rising markets will also lead to higher oil demand, inducing an upward movement in oil prices. As financialized commodities, such as gold, respond faster than goods markets, this may generate predictive power from gold to oil. However, the most significant causality from gold to oil should be observed when investors shift large amounts of funds between the gold and oil markets. Such shifts occur during large financial crises.

Few studies examine oil and gold prices and their relationship with macroeconomic and financial variables in the literature. However, researchers were more interested in studying the interrelations between oil and gold following the new economic crisis due to the increases in the prices of these commodities and the common use of gold as a safe haven for their investments. The pioneering study in this respect was the one conducted by Melvin and Sultan (1990), in which the researchers found a strong correlation between oil and gold due to the export revenue channel. Another study was performed by Kim and Dilts (2011), where they had similar observations. Other studies, including Soytas, Sari, Hammoudeh, and Hacıhasanoglu (2009), Liao and Chen (2008), Sari, Hammoudeh, and Ewing (2007), Hammoudeh and Yuan (2008), Narayan et al. (2010), Šimáková (2011), Le and Chang (2011b), and Lee, Huang, and Yang (2012), however, did not find evidence of the relationship between the movement of the prices of oil and gold. Shahbaz, Balcilar, and Ozdemir (2017), using a nonparametric causality-in-quantiles test, showed that the oil price has weak predictive power for the gold price, and the causality-in-variance tests found strong support for the predictive capacity of oil for gold market volatility. According to others, the prices of oil and gold act simultaneously because they are correlated with the movement of their long-term driving factors, such as volatility in U.S. dollars and the turmoil in the international politics (e.g., Bampinas & Panagiotidis, 2015; Le & Chang, 2011a). This simultaneous movement can generate dynamic causality links between the oil and gold markets.

With respect to literature's main conclusions concerning on the investigation of the dynamic nexus between gold market and oil market, the existing studies employing the methods fall into two groups: those that use a linear Granger causality and non-linear Granger causality methods and those that use causality-in-quantile method. First group studies examining the causality between gold market and oil market using these methods have been conducted by Zhang and Wei (2010), Jain and Ghosh (2013), Bildirici and Turkmen (2015), Jain and Biswal (2016), Kumar (2017), Kanjilal and Ghosh (2017), Gil-Alana, Yaya, and Awe (2017), Bilgin, Gogolin, Lau, and Vigne (2018), Sephton and Mann (2018), and Mei-Se, Shu-Jung, and Chien-Chiang (2018). For instance, Bildirici and Turkmen (2015), Jain and Biswal (2016), Kumar (2017), Kanjilal and Ghosh (2017), Mei-Se et al. (2018), and Sephton and Mann (2018) show that there is a bidirectional causality between crude oil prices and gold prices. But Zhang and Wei (2010) indicate that the crude oil price change linearly Granger causes the volatility of gold price, but not vice versa. Second group studies investigating the causality between both markets using causality-in-quantile method have been showed by Shahbaz et al. (2017) and Das, Kumar, Tiwari, Shahbaz, and Hasim (2018). Shahbaz et al. show that oil price has a weak predictive power for the gold price using the nonparametric causality-in-quantiles test. Furthermore, the study indicates strong support for the predictive capacity of oil for gold market volatility. Das et al. examined the dependence structure of stocks, gold, and crude oil with financial stress using the nonparametric causality-in-quantile method. The evidence of study shows that there is an evidence of bilateral causality in mean and variance for gold and crude oil with respect to financial stress and stocks to be influential to financial stress in both mean and variance.

The findings obtained from the above studies do significantly vary, and the variation in findings can be attributed to different country case studies, different data samples, and employing different types of estimation approaches. Economic causality has tended to rely on justifications from economic theory to infer the direction of dynamic links between variables and to inform empirical testing of causal hypotheses. However, there are no relevant theoretical bases most of the time in determining the empirical relationships between variables that appear jointly determined over time. Thus, there are difficulties in interpretation, test execution, and handling additional relevant variables. This is also the case for studies examining the oil–gold nexus. Because no clear theoretical economic model exists on the relationship between oil and gold, previous studies often employed Granger causality to investigate the dynamic links between oil and gold prices. Granger causality was popular in those studies partly because it was not specific to a particular structural model but instead depended on the stochastic nature of variables. A Granger causality test may be sensitive to the time period of estimation (Arora & Shi, 2015; Balcilar & Ozdemir, 2013a; Balcilar, Ozdemir, & Arslanturk, 2010; Hurn, Phillips, & Shi, 2016; Psaradakis, Ravn, & Sola, 2005; Shi, Hurn, & Phillips, 2016; Stern, 2000; Stock & Watson, 1989; Swanson, 1998; Thoma, 1994), which is also the case for the studies examining the Granger causality between the oil and gold prices. In the literature, various methods have been used to accommodate the time-varying nature of the causal link between series. One is a forward expanding window version of the Granger causality test (Swanson, 1998; Thoma, 1994), and the others include a rolling Granger causality test (Arora & Shi, 2015; Balcilar et al., 2010; Balcilar & Ozdemir, 2013a; Swanson, 1998), recursive rolling Granger causality test (Hurn et al., 2016; Shi et al., 2016), and a Markov-switching Granger causality test (Psaradakis et al., 2005). However, none treats the relationship between oil and gold prices comprehensively because both the oil and gold markets are subject to frequent structural breaks. Both oil and gold are globally traded commodities, and the markets for these commodities are subject to a number of events, such as wars, demand hikes, financial crises, policy changes, technological innovations, and political events, which are likely to induce shifts in the dynamic links between oil and gold prices. Previous studies on the dynamic links between oil and gold prices all assume constant parameters; therefore, their findings might be invalid.

Structural breaks (or regime shifts) are one of the most challenging issues for time series econometric methods (Granger, 1996). Hansen (2001) and Perron (2006) affirm that econometric application estimations involving time series data should distinctly consider the effects of structural breaks or regime shifts. In the presence of structural breaks, the parameters of the econometric models show time variation, and statistical tests based on the constant parameter assumption, such as the Granger causality, are invalid and lead to incorrect inferences. In this paper, we perform time-varying Granger causality tests between oil and gold price series.

To fill the gap in the relevant literature on the unfamiliarity of the effect of structural breaks on the oil–gold relationship, we use a time-varying Granger causality test from Balcilar et al. (2010) and Balcilar and Ozdemir (2013a) and the recursive rolling test from Hurn et al. (2016) and Shi et al. (2016) to examine the time-varying nature of the causal nexus of oil and gold for G-7 countries. We use monthly data for the spot and futures prices of the oil and gold markets ranging between 1- and 12-month maturities. Because oil market and gold market began operations at different dates, the data span is different for each series of both markets. To the best of our knowledge, this is the first paper that analyses the time-varying causal nexus between oil and gold using the rolling Granger causality test of Balcilar et al. and Balcilar and Ozdemir and the recursive rolling Granger causality test of Hurn et al. and Shi et al.

As we have discussed above, the previous literature can be primarily grouped into linear or non-linear cointegration-causality type studies and those studies considering causality-in-quantiles approach. All the linear or non-linear cointegration-causality type studies are based on full sample analysis and asymptotic statistical inference, except Mei-Se et al. (2018). Mei-Se et al. examines the time-varying cointegration relationship among three metals, including gold and oil. Hurn et al. (2016) and Shi et al. (2016) find that based on Monte Carlo simulations, recursive approach to exhibit least performance with reference to both true and false causality detection rates. Moreover, the cointegration only implies existence of an at least way causality link among series (Engle & Granger, 1987) but does not discriminate other type of causalities. As we discussed in the next section, all of these asymptotic testing approaches might have series size and power distortions in finite samples and not robust to integration–cointegration properties of the data (Balcilar et al., 2010; Balcilar & Ozdemir, 2013a; Dolado

& Lütkepohl, 1996; Park & Phillips, 1989; Sims, Stock, & Watson, 1990; Toda & Phillips, 1993; Toda & Phillips, 1994; Toda & Yamamoto, 1995; and Yamada & Toda, 1998). The causality-in-quantiles tests are robust to misspecification and structural breaks (Balcilar, Bekiros, & Gupta, 2017a; Balcilar, Gupta, & Pierdzioch, 2017b). Although, both the nonparametric causality-in-quantiles tests and rolling and recursive rolling tests used in this study are robust against structural breaks, each one has certain advantages. The causality-in-quantiles approach is a full sample method. Although, it can consider causality in various quantiles, it lacks the time perspective and will not uncover periods when causal links exist and when they do not. Rolling and recursive rolling approaches used in this study are not full sample methods and does not consider a certain quantile over the whole sample period. They rather look through time and identify how causality links evolve over the time window. Thus, by using rolling and recursive rolling approaches, one can detect the periods where causality actually may or may not exists. On the other hand, it is not possible to date-stamp periods where causality exists using the causality-in-quantiles approach. Rolling and recursive rolling approaches also allow us to study the effects of particular events on the causality links. It is not actually possible to link a particular event to causal or noncausal periods using the causality-in-quantiles approach. Thus, rolling and causality-in-quantiles approaches, although both are robust methods, provide different information on the causality links.

The findings obtained from the rolling and the recursive rolling approaches show that the causality relationship between oil and gold prices in both the spot and futures markets displays significant time variation with several switches from bidirectional causality to unidirectional or noncausality. We find several periods of bidirectional causal links between oil prices and gold prices over the 1983–2016 sample period for spot market and futures market at 1- to 12-month maturities. However, there are longer periods of noncausality links over the study period. Indeed, noncausality is more prevalent than causality.

The rest of this study is organized as follows: The next section provides the methodology; Section 3 evaluates the data and empirical findings; and the last section presents the conclusion.

2. Methodology

To test for Granger causality in the presence of structural breaks, one of the approaches used in the literature is regime-switching models, such as the Markov-switching model (Hamilton, 1989; Krolzig, 1999) and threshold autoregression (Granger and Teräsvirta, 1993; Deutsch, Granger, & Teräsvirta, 1994; Teräsvirta, 1998). Balcilar and Ozdemir (2013b, 2013c) used two-regime Markov-switching vector autoregressive (VAR) models for a causality analysis in the presence of Markov switching. Although regime-switching models can be used for a regime-switching Granger causality analysis, they are restricted in the sense that there are usually two or three regimes, and they will not be able to capture multiple regime changes with varying degrees of regime duration. As an alternative to regime-switching models, the recursive estimation approach of Swanson (1998) and Thoma (1994); the rolling window estimation approach of Balcilar et al. (2010), Balcilar and Ozdemir (2013a), and Swanson (1998); and the recursive rolling estimation approach of Hurn et al. (2016) and Shi et al. (2016) can be used to conduct time-varying Granger causality tests. Each of these approaches have their strengths and weaknesses, but they all allow for multiple structural breaks with possible shifts in parameters in each time period. For time-varying Granger causality tests where the systematic shifts between two or three regimes is not the dominant feature but parameters are subject to many shift, the recursive, rolling, or recursive rolling approaches should be chosen (see Balcilar et al., 2010; Hurn et al., 2016). In this study, we prefer the rolling and recursive rolling approaches due to their strengths in terms of false and true causality detection rates. Hurn et al. and Shi et al. perform Monte Carlo simulations to compare the false and successful causality detection rates of the recursive, rolling, and recursive rolling methods. They find that the recursive approach has the worst performance in terms of both false and successful detection rates. The rolling approach using the implementation in Balcilar et al. and Balcilar and Ozdemir has the highest successful detection rate, whereas the recursive rolling approach has a slightly lower false detection rate compared with the rolling approach. Shi et al. find that the rolling approach has the best overall performance for integrated time series.

To explain the time-varying causality tests conducted for the causality links between oil prices (y_{1t}) and gold prices (y_{2t}), consider the following bivariate VAR(p):

$$y_{1t} = \phi_{10} + \sum_{i=1}^{p} \phi_{11,i} y_{1t-i} + \sum_{i=1}^{p} \phi_{12,i} y_{2t-i} + \varepsilon_{1t}, \qquad (1)$$

$$y_{2t} = \phi_{20} + \sum_{i=1}^{p} \phi_{21,i} y_{1t-i} + \sum_{i=1}^{p} \phi_{22,i} y_{2t-i} + \varepsilon_{2t}, \qquad (2)$$

where p is the lag order, and ε_{it}, $i = 1, 2$ are white noise error terms. Let us write this VAR(p) in the following matrix notation:

$$y_t = \Phi_0 + \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \ldots + \Phi_p y_{t-p} + \varepsilon_t, \quad (3)$$

with following companion multivariate form:

$$y_t = \Pi x_t + \varepsilon_t, \qquad t = 1,\ 2,\ \ldots,\ T, \quad (4)$$

where $y_t = (y_{1t}, y_{2t})', x_t = \left(1, y'_{t-1}, y'_{t-2}, \ldots, y'_{-p}\right)', \varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$, and $\Pi = [\Phi_0, \Phi_1, \ldots, \Phi_p]$ is a $2 \times (2p + 1)$ matrix.

The Granger noncausality from the gold price to oil price implies that gold does not have predictive power for oil ($y_{2t} \nrightarrow y_{1t}$), whereas the Granger noncausality from the oil price to the gold price implies that oil does not have predictive power for gold ($y_{1t} \nrightarrow y_{2t}$). The two noncausality statements, respectively, lead to the following joint restrictions to be tested under the following null hypotheses:

$$H_0: y_{2t} \nrightarrow y_{1t} \Longrightarrow \phi_{12,1} = \phi_{12,2} = \ldots = \phi_{12,p} = 0, \quad (5)$$

$$H_0: y_{1t} \nrightarrow y_{2t} \Longrightarrow \phi_{21,1} = \phi_{21,2} = \ldots = \phi_{21,p} = 0. \quad (6)$$

The Granger noncausality restrictions in Equations 5 and 6 can be written concisely as

$$H_0: R\pi = 0, \quad (7)$$

where $\pi = \mathrm{vec}(\Pi)$ is a vector of dimension $2(2p + 1) \times 1$ using row vectorization and R is the selection matrix of dimension $p \times 2(2p + 1)$.

The rolling estimation approach uses a fixed window size $\tau_w = [r_w T]$, where $[\cdot]$ is the integer part with fraction r_w of the total number of observations. The estimation starting point is $\tau_1 = [r_1 T]$ with fraction r_1, and the end point is $\tau_2 = [r_2 T]$ with fraction r_2, giving a sequence of Wald statistics for $t = \tau_1, \tau_1 + 1, \ldots, \tau_2$. The rolling procedure can be represented with start point $\tau_1 = \tau_2 - \tau_w + 1$ and the sequence of end points $\tau_2 = \{\tau_w, \ldots, T\}$. In this manner, we obtain the $T - \tau_w + 1$ sequence of Wald statistics. We denote the sequence of the rolling statistics with $\left\{W^{\tau_2}_{\tau_1 = \tau_2 - \tau_w + 1}\right\}_{\tau_2 \in [\tau_w, T]}$.

The recursive rolling approach is similar to the rolling approach and likewise is a fixed window estimation method with window size τ_w. Similar to the rolling estimation, the end point of the regression is the sequence $\tau_2 = \{\tau_w, \ldots, T\}$. However, the start point of the estimation considers all possibilities, which is 1 to $\tau_2 - \tau_w + 1$. The recursive rolling procedure combines the sequence of end points $\tau_2 = \{\tau_w, \ldots, T\}$ with the start point sequence $\tau_1 = \{1, \tau_2 - \tau_w + 1\}$. The recursive rolling statistics are the sum of the all possible rolling statistics for the given point and are denoted with $\left\{SW^{\tau_2}_{\tau_{1=\tau_2-\tau_w+1}}\right\}_{\tau_2 \in [\tau_w, T]} = \sup_{\tau_2, \tau_1^s \in [1, \tau_2 - \tau_w + 1]} \left[\left\{W^{\tau_2}_{\tau_1^s}\right\}\right]$.

In the practical implementation of the rolling and recursive rolling approaches, we need to calculate Wald test sequences for a subset of the sample with start point τ_1 and end point τ_2. Let the ordinary least squares (OLS) estimate of the VAR(p) model in Equation (4) estimated for this subsample be given by $\widehat{\Pi}_{\tau_1,\tau_2}$ and its row vec form by $\widehat{\pi}_{\tau_1,\tau_2} = \mathrm{vec}\left(\widehat{\Pi}_{\tau_1,\tau_2}\right)$. The sequence of Wald tests is obtained by imposing the restrictions in Equation (7) on the subsample estimates; that is, we test the null hypothesis with restrictions $R\widehat{\pi}_{\tau_1,\tau_2} = 0$. The OLS estimates $\widehat{\pi}_{\tau_1,\tau_2}$ are given for each equation $i = 1, 2$ by $\widehat{\pi}_{i,\tau_1,\tau_2} = \left[\sum_{t=\tau_1}^{\tau_2} y_{it}\ x'_t\right]\left[\sum_{t=\tau_1}^{\tau_2} x_t\ x'_t\right]^{-1}$. The residuals for each equation in the subset estimate can be obtained as $\widehat{\varepsilon}'_t = [\widehat{\varepsilon}_{1t}, \widehat{\varepsilon}_{1t}]$ with $\widehat{\varepsilon}_{it} = y_{it} - \widehat{\pi}_{i,\tau_1,\tau_2} x_t$. We can obtain the corresponding estimate of the residual covariance matrix Ω as $\widehat{\Omega}_{\tau_1,\tau_2} = T_w^{-1} \sum_{t=\tau_1}^{\tau_2} \widehat{\varepsilon}_t\ \widehat{\varepsilon}'_t$, where $T_w = \tau_2 - \tau_1 + 1$. Given these definitions, the Wald statistics for the Granger noncausality restrictions for each subsample can be obtained from

$$W^{\tau_2}_{\tau_1} = \left(R\widehat{\pi}_{\tau_1,\tau_2}\right)' \left\{R\left[\widehat{\Omega}_{\tau_1,\tau_2} \otimes \left(\sum_{t=\tau_1}^{\tau_2} x_t x'_t\right)^{-1}\right] R'\right\}^{-1} \left(R\widehat{\pi}_{\tau_1,\tau_2}\right). \quad (8)$$

The Wald test statistics in Equation (8) assume that homoskedastic errors and Granger causality tests constructed in this way may have invalid empirical levels and could be accompanied by power loss when the errors are heteroskedastic. To avoid issues due to unconditional and conditional heteroskedasticity, we also use a modified Wald test that accounts for the effects of heteroskedasticity in the residuals. The modified Wald test is constructed as follows:

$$W^{*\tau_2}_{\tau_1} = T_w \left(R\widehat{\pi}_{\tau_1,\tau_2}\right)' \left[R\left(\widehat{V}^{-1}_{\tau_1,\tau_2} \widehat{W}_{\tau_1,\tau_2} \widehat{V}^{-1}_{\tau_1,\tau_2}\right) R'\right]^{-1} \left(R\widehat{\pi}_{\tau_1,\tau_2}\right), \quad (9)$$

where $\widehat{V}_{\tau_1,\tau_2} = I_2 \otimes \widehat{Q}_{\tau_1,\tau_2}$, $\widehat{Q}_{\tau_1,\tau_2} = T_w^{-1} \sum_{t=\tau_1}^{\tau_2} x_t\, x'_t$, and $\widehat{W}_{\tau_1,\tau_2} = T_w^{-1} \sum_{t=\tau_1}^{\tau_2} \xi_t\, \xi'_t$ with $\xi_t = \widehat{\varepsilon}_t \otimes x_t$.

The asymptotic distribution of the rolling Wald test is a chi-square, and the asymptotic distribution of the recursive rolling statistic is non-standard but is given in Hurn et al. (2016). There is significant evidence (Guilkey & Salemi, 1982; Toda & Phillips, 1993; Toda & Phillips, 1994) that the Wald tests, including the Granger causality test used in this study, may suffer from serious size distortions. Moreover, the modified Wald test in Equation (9) involves the estimation of the matrix $\widehat{W}_{\tau_1,\tau_2}$, which is essentially a matrix of the fourth moment. A fourth moment estimator will be much more sensitive to high variations in small samples. We therefore follow the bootstrap approach of Balcilar et al. (2010) to obtain the empirical distributions of the Wald tests.[1,2] The bootstrap procedure is implemented by considering the fact that under the Granger noncausality null (maintained) hypothesis restrictions $R\widehat{\pi}_{\tau_1,\tau_2} = 0$, the VAR($p$) model has constant coefficients $\Pi_{\tau_1,\tau_2} = \Pi$ for all (fractional) subsamples $t = \tau_1, \tau_1 + 1, ..., \tau_2$.

One important issue in Granger causality testing is the integration–cointegration properties of the data. As shown by Sims et al. (1990), Park and Phillips (1989), and Toda and Phillips (1993, 1994), the Wald tests of Granger noncausality in Equations 8 and 9 based on levels estimation do not usually have standard asymptotic distribution. Moreover, these Wald tests will also depend on the nuisance parameters when data have unit roots, that is, integrated of order one, denoted I(I). Sims et al. showed that one cannot perform Granger causality tests by imposing the joint restrictions in Equations 8 and 9 based on standard asymptotic theory using the OLS estimation of the underlying VAR model in Equations 1 and 2. In the case of I(1) data, there are three alternative approaches proposed in the literature for preforming Granger noncausality tests in Equations 5 and 6. First approach is proposed by Mosconi and Giannini (1992) and Toda and Phillips (1993) and utilizes the version of the error correction model (ECM) in Johansen (1988, 1991). In this approach, equivalent restrictions to those in Equations 5 and 6 on the original VAR model are imposed on the ECM. The second approach is the fully modified VAR (FM-VAR) approach of Phillips (1995). Analogous to Johansen-type ECM approach, one can impose equivalent Granger noncausality restrictions on the FM-VAR. The third approach is the lag-augmented VAR (LA-VAR) approach (Dolado & Lütkepohl, 1996; Toda & Yamamoto, 1995; Yamada & Toda, 1998). In the LA-VAR approach, the lag order p of the VAR is extended as $p + d_{\max}$, where $d_{\max}$ is the maximum integration order of the series, and Granger causality restrictions in Equations 5 and 6 are imposed on the first p lags. In all three approaches, the Wald test statistics relating to Equations 5 and 6 follow standard chi-squared distribution. In terms of performance, Monte Carlo studies performed by Yamada and Toda (1998) show that the LA-VAR test has better size and stability properties compared with the FM-VAR and the ECM approaches. In this paper, we do not use any of these approaches due to their drawbacks. In the ECM approach, testing the Granger noncausality hypothesis involves non-linear restrictions on the parameter restrictions; and therefore, Wald or likelihood ratio tests suffer from size distortions due to the rank deficiency problem (Toda & Phillips, 1993). Another problem with the ECM approach is the requirement of the prior knowledge of cointegrations rank, which leads to the size distortions due to the pretest bias (Yamada & Toda, 1998). As pointed out by Yamada and Toda (1998), LA-VAR approach uses sample information inefficiently, which leads to size and power distortions in finite samples. This may be a serious issue in our case because the rolling and recursive rolling approaches have relatively smaller sample sizes. The FM-VAR approach does not always achieve good asymptotic size and additionally, depending on the location and number of unit roots in the system, it may be quite conservative under the null hypothesis, leading to low power under the alternative (Yamada & Toda, 1998).

The rolling and recursive rolling Granger causality tests in our case involves relatively smaller sample sizes. Therefore, the ECM-, FM-VAR-, and LA-VAR-based Granger noncausality tests are very likely to perform poorly. In order avoid the size and power distortion issues in the asymptotic test approaches, we use bootstrap Wald test approach of Balcilar et al. (2010) and Balcilar and Ozdemir (2013a). Pioneered by Efron (1979), the bootstrap method uses critical or p values generated from the empirical distribution of the Wald tests derived from the sample data under the null hypotheses in Equations 5 and 6. The bootstrap approach for Granger noncausality testing obtains valid critical or p values irrespective of the integration–cointegration properties of the data because these are obtained from the empirical distribution derived from the sample data. Horowitz (1994), Mantalos and Shukur (1998), and Mantalos (2000), among others, documented the robustness of the bootstrap approach for testing Granger noncausality. The Monte Carlo simulations performed by Hacker and Hatemi-J (2006) showed that the bootstrap Wald test has much smaller size distortions compared with the use of asymptotic tests. On the basis of Monte Carlo simulations, Mantalos and Shukur and Mantalos showed that these results hold irrespective of sample sizes, integration orders, and error-correction processes (homoscedastic or autoregressive conditional heteroskedasticity [ARCH]). Therefore, we adopt the bootstrap approach for Granger noncausality testing because of its advantages. The critical values (or the p values) of the Granger

[1]The details of the bootstrap implementation can be found in Balcilar et al. (2010).

[2]For integrated time series data, rolling, and recursive rolling Granger causality Wald tests can be performed using the LA-VAR approach of Toda and Yamamoto (1995). Shi et al. (2016) adopt an LA-VAR approach to perform rolling and recursive rolling Granger causality Wald tests. In this study, we use bootstrap approach in order obtain the critical or p values of the Wald statistics, which is robust to integration–cointegration properties of the data (Mantalos, 2000; Mantalos & Shukur, 1998).

causality tests are obtained under the restrictions of the null using the residual-based bootstrap approach of Balcilar et al. with 1,000 replications.

3. Data and empirical findings

In this paper, we use the monthly U.S. dollar closing prices of nearby settlement crude oil spot and futures contracts traded on the New York Mercantile Exchange and the monthly U.S. dollar prices of gold spot and futures contracts traded on the London Bullion Market. Data on gold and oil spots and futures prices are obtained from DataStream, and we use the data of the gold and oil markets for spot and futures prices at 1- to 12-month maturities. The last column of Table 1 shows the oil series and gold series included in the sample, along with the beginning date of each series. Accordingly, the data span is from the date of reporting to August 2016. Time series plots of the spot and futures prices of oil and gold are shown in Figure 1.

The theory, as well as empirical evidence, indicates that different markets are likely to have differing degrees of sensitivity to oil prices. A price on the gold market, for instance, would have an averaging effect across markets and would not perhaps reveal the links between oil and gold prices. We employ monthly frequency data rather than daily frequency data because daily data may be affected by drifts and noise, which probably serve to mask the dependence relationship and complicate the modelling of the marginal distributions in the presence of non-stationary variances, long memory, or sudden jumps; it would also be difficult to capture links between oil prices and gold prices on a daily basis, as high frequency oil and gold prices are highly volatile. Oil and gold prices are both nonstationary at log levels, as standard unit root tests[3] suggest. Given that all series are nonstationary, we additionally check for cointegration between the pairs of spot and 1- to 12-month futures of oil and gold price series. The Johansen (1988, 1991) cointegration tests indicate no cointegration for each pair of the series.[4] As we have explained in the previous section, bootstrap Granger causality tests used in the study are valid irrespective of the integration–cointegration properties of the data, so no special treatment of the unit roots and non-existence of cointegration are needed.

Table 1 presents selected descriptive key features of the data series, the mean, standard deviation, kurtosis, skewness, the Jarque–Bera normality test, the Ljung–Box first Q([1]) and the fifth Q([5]) autocorrelation tests, and the first- [ARCH(1)] and the fifth- [ARCH(5)] order Lagrange multiplier tests for the ARCH for oil and gold spots and futures prices series in natural logarithms. As is clear from the table, the mean for oil market prices is the lowest for the futures price, but it increases slightly for spot prices, with the longest maturity of a 12-month contract having the lowest average price of approximately 3.49. The mean of gold market prices is highest at the 12-month maturity contract for the futures price but then slightly decreases to the spot price and maturity of 1 month for the futures price contract with an average price of 6.23. We also observe from the table that oil markets are more volatile than gold markets. The positive values of the skewness statistic suggest a lower probability of large decreases in prices. The return distributions have thin tails, as indicated by negative values for the kurtosis statistic. Both variables are skewed to the right with negative kurtosis, resulting in non-normal distributions. The values of the Ljung–Box statistic show that there is serial correlation in oil market prices and the gold market prices. Finally, the autoregressive conditional heteroskedasticity–Lagrange multiplier statistics indicate that ARCH effects exist in all price series.

Even though we aim to analyse the time-varying nature of the causal nexus between oil and gold, we also performed the linear Granger causality test based on a full sample VAR model for completeness and comparability.[5] The results of the linear Granger causality tests are reported in Table 2. The second column of Table 2 shows the linear Granger causality test results for testing the null hypothesis that oil prices do not Granger cause gold prices, whereas the third column of Table 2 reports the linear Granger causality test results for testing the null hypothesis that gold prices do not Granger cause oil prices. The evidence from column 2 of Table 2 indicates that the null hypothesis that oil prices do not Granger cause gold prices is not rejected at the 5% level of significance. The F-test results from column 3 of Table 2 do not reject the null hypothesis that gold prices do not Granger cause oil prices at a 5% level of significance. In sum, the results of the linear Granger causality test show that there is no evidence of predictability emanating between oil and gold markets at the 5% level of significance.

The results reported in Table 2 for examining the causal link between oil and gold prices in spot and futures markets are based on restrictions imposed in a

[3]Complete details of the unit root tests are available upon request from the authors. See also Shahbaz et al. (2017).

[4]Details of the cointegration tests are not reported to preserve space but available from the authors upon request.

[5]Due to the nonstationarity of all series, all full sample Granger causality tests are performed using the LA-VAR approach of Toda and Yamamoto (1995).

Table 1. Descriptive statistics for the oil and gold price series in natural logarithms

	n	Mean	*SD*	Min	Max	Skewness	Kurtosis	JB	Q(1)	Q(4)	ARCH(1)	ARCH(4)	Starting period
	Panel A: Oil price series												
Spot	368	3.53	0.67	2.42	4.98	0.41	−1.27	34.42***	362.69***	1,390.31***	326.45***	327.48***	Dec. 1985–Aug. 2016
1 month	401	3.51	0.65	2.42	4.98	0.48	−1.09	35.13***	395.01***	1,516.86***	361.59***	359.80***	Apr. 1983–Aug. 2016
2 months	401	3.51	0.65	2.43	4.98	0.49	−1.13	36.69***	396.04***	1,528.10***	363.68***	361.81***	Apr. 1983–Aug. 2016
3 months	401	3.51	0.65	2.44	4.98	0.49	−1.15	38.16***	396.88***	1,536.87***	365.84***	364.03***	Apr. 1983–Aug. 2016
4 months	401	3.51	0.66	2.45	4.99	0.50	−1.17	39.52***	397.55***	1,543.97***	367.65***	365.91***	Apr. 1983-Aug. 2016
5 months	401	3.51	0.66	2.46	4.99	0.51	−1.19	40.77***	398.08***	1,549.79***	369.12***	367.42***	Apr. 1983–Aug. 2016
6 months	401	3.50	0.66	2.46	4.99	0.52	−1.20	41.84***	398.49***	1,554.54***	370.33***	368.65***	Apr. 1983–Aug. 2016
7 months	401	3.50	0.66	2.47	4.99	0.52	−1.21	42.77***	398.85***	1,558.47***	371.41***	369.76***	Apr. 1983–Aug. 2016
8 months	393	3.50	0.67	2.48	4.99	0.52	−1.26	43.45***	391.28***	1,531.16***	364.09***	362.47***	Nov. 1983–Aug.2016
9 months	393	3.50	0.67	2.48	4.99	0.52	−1.27	44.16***	391.52***	1,533.97***	364.94***	363.32***	Nov. 1983–Aug.2016
10 months	393	3.50	0.67	2.49	4.99	0.53	−1.28	44.80***	391.73***	1,536.46***	365.66***	364.05***	Nov. 1983–Aug.2016
11 months	393	3.50	0.67	2.50	4.99	0.53	−1.29	45.19***	391.84***	1,538.38***	365.96***	364.16***	Nov. 1983–Aug.2016
12 months	392	3.49	0.68	2.46	4.98	0.51	−1.29	43.76***	390.78***	1,533.86***	365.13***	363.50***	Dec. 1983–Aug. 2016
	Panel B: Gold price series												
Spot	401	6.23	0.57	5.54	7.54	0.90	−0.68	61.56***	398.44***	1,571.60***	390.63***	388.06***	Apr. 1983–Aug. 2016
1 month	401	6.23	0.57	5.54	7.54	0.90	−0.68	61.56***	398.44***	1,571.58***	390.79***	388.22***	Apr. 1983–Aug. 2016
2 months	401	6.24	0.57	5.54	7.54	0.89	−0.68	61.40***	398.43***	1,571.44***	390.74***	388.17***	Apr. 1983–Aug. 2016
3 months	401	6.24	0.57	5.54	7.54	0.89	−0.68	61.14***	398.41***	1,571.29***	390.66***	388.09***	Apr. 1983-Aug. 2016
4 month	401	6.24	0.56	5.54	7.54	0.89	−0.68	60.70***	398.34***	1,570.99***	390.54***	387.97***	Apr. 1983–Aug. 2016
5 months	401	6.25	0.56	5.55	7.54	0.88	−0.69	60.02***	398.31***	1,570.62***	390.44***	387.86***	Apr. 1983–Aug. 2016
6 months	401	6.26	0.56	5.55	7.54	0.88	−0.69	59.44***	398.31***	1,570.75***	390.37***	387.79***	Apr. 1983–Aug. 2016
7 month	401	6.26	0.56	5.56	7.54	0.87	−0.69	58.75***	398.30***	1,570.65***	390.30***	387.72***	Apr. 1983–Aug. 2016
8 months	401	6.27	0.55	5.56	7.54	0.86	−0.70	58.02***	398.28***	1,570.55***	390.23***	387.65***	Apr. 1983–Aug. 2016
9 months	401	6.28	0.55	5.57	7.54	0.86	−0.70	57.28***	398.27***	1,570.45***	390.16***	387.59***	Apr. 1983–Aug. 2016
10 months	401	6.28	0.55	5.57	7.54	0.85	−0.70	56.51***	398.25***	1,570.34***	390.10***	387.52***	Apr. 1983–Aug. 2016
11 months	401	6.29	0.55	5.58	7.54	0.84	−0.71	55.71***	398.24***	1,570.24***	390.03***	387.45***	Apr. 1983–Aug. 2016
12 months	401	6.30	0.54	5.59	7.55	0.83	−0.71	54.89***	398.23***	1,570.15***	389.97***	387.39***	Apr. 1983–Aug. 2016

Note. ARCH: autoregressive conditional heteroskedasticity; JB: Jarque–Bera normality test; *SD*: standard deviation. The table reports the descriptive statistics for the spot and futures (1 to 12 month) prices for the oil (Panel A) and gold (Panel B) markets. Data have *n* observations and are reported monthly with the sample periods in the last column of the table. In addition to the mean, the *SD*, minimum (min), maximum (max), skewness, and kurtosis statistics, the table reports the JB, the Ljung–Box first Q([1]) and the fourth Q([4]) autocorrelation tests, and the first- [ARCH(1)] and the fourth- [ARCH(4)] order Lagrange multiplier tests for the ARCH.

*Significance at 10% level. **Significance at 5% level. ***Significance at 1% level.

Note: Figure plots of the natural logarithms of the price series.

Figure 1. Time series plots of the spot and futures prices for the oil and gold markets

Table 2. Linear Granger causality tests

	H_0: Oil prices do not Granger cause gold prices	H_0: Gold prices do not Granger cause oil prices	Order of the VAR (*p*)
Spot	0.5423	0.6567	1
1 month	0.4049	0.5325	1
2 months	0.5508	0.6378	1
3 months	0.6186	0.7961	1
4 months	0.6535	1.0948	1
5 months	0.5251	1.4874	1
6 months	0.6844	0.9659	1
7 months	0.5917	0.8455	1
8 months	0.5951	0.7968	1
9 months	0.6388	0.723	1
10 months	0.501	0.5093	1
11 months	1.4843	0.3241	1
12 months	0.6522	0.4597	1

Note. VAR: vector autoregressive. The table reports the F statistic for the no-Granger causality restrictions imposed on a lag-augmented linear VAR model under the null hypotheses H_0. The order (p) of the VAR is selected by the Bayesian Information Criterion. ***, **, and * indicate rejection of the null of no Granger causality at a 1%, 5%, and 10% level of significance, respectively.

linear VAR model estimated for the full sample. This full sample VAR model assumes that no structural breaks exist in the sample and that parameters are constant over the entire sample period. However, structural changes may shift the parameter values, and the patterns of causal relationships may vary over time. That is, structural changes may influence the temporal (Granger) causality links, being sensitive to the sample period adopted. Several tests are available to investigate the stability of VAR models (Andrews & Ploberger, 1994). When the estimated parameters come from unstable relationships that are undetected, problems are likely to arise. It is clear that such parameter estimates stemming from unstable relationships can lead to serious consequences (Hansen, 1992) due to biased inferences as well as inaccurate forecasts (Zeileis, Leisch, Hornik, & Kleiber, 2005). We test the stability of the VAR model parameters to examine the temporal stability of the coefficients of the VAR model composed of oil and gold prices. We employ three different statistics (*Sup-F*, *Mean-F*, and *Exp-F*) suggested by Andrews (1993) and Andrews and Ploberger (1994). All three tests require trimming from the ends of the sample for the stability of short-term parameters. Table 3 shows the results of the parameter stability test performed for oil and gold prices. We derive the critical values and the *p* values using the parametric bootstrap

Table 3. Parameter stability tests

	Oil equation			Gold equation			VAR system		
	Sup-F	***Mean-F***	***Exp-F***	***Sup-F***	***Mean-F***	***Exp-F***	***Sup-F***	**Mean-F**	***Exp-F***
Spot	16.499**	6.643**	5.352**	17.741***	8.259**	5.682**	23.994**	14.274***	8.361**
1 month	16.996**	6.064*	5.212**	22.330***	13.026***	8.291***	25.264***	13.003***	8.303**
2 months	16.807**	5.700*	5.079**	21.958***	13.108***	8.143***	26.081***	13.056***	8.385**
3 months	17.193**	5.564*	5.070**	21.592***	13.072***	7.996***	27.217***	13.008***	8.620**
4 months	17.466**	5.567*	5.085**	21.756***	13.457***	8.151***	28.284***	13.158***	8.955***
5 month	17.724**	5.629*	5.132**	21.033***	13.213***	7.937***	29.422***	13.171***	9.371***
6 months	17.995***	5.751*	5.209**	20.575***	12.807***	7.715***	30.463***	13.080***	9.796***
7 months	18.137***	5.749*	5.219**	20.161***	12.586***	7.574***	31.451***	13.067***	10.229***
8 months	18.802***	6.124**	5.478**	18.666***	11.184***	6.738***	31.870***	13.608***	10.442***
9 months	18.979***	6.173**	5.533**	18.364***	10.957***	6.611***	32.514***	13.551***	10.734***
10 months	19.218***	6.319**	5.631**	18.176***	11.001***	6.633***	32.972***	13.693***	10.949***
11 months	19.400***	6.579**	5.794**	17.667**	10.152***	6.150**	32.685***	13.485***	10.805***
12 months	16.666**	5.446*	4.668**	16.773**	8.594***	5.314**	31.144***	12.747**	10.060***

Note. VAR: vector autoregressive. The parameter stability tests exhibit non-standard asymptotic distributions. With the parametric bootstrap procedure, Andrews (1993) and Andrews and Ploberger (1994) report the critical values and *p* values for the non-standard asymptotic distributions of these tests. Additionally, according to Andrews (1993), trimming from both ends of the sample is required for the *Sup-F, Mean-F*, and *Exp-F*. Hence, the tests are applied to the fraction of the sample in (0.15, 0.85), that is, a 15% trimming from each end of the sample. We calculate the critical values of the tests using 2,000 bootstrap replications.

***, **, and * indicate significance at a 1%, 5%, and 10% level, respectively.

distribution obtained using 2,000 replications generated from a VAR model with constant parameters (Andrews, 1993). A 15% trimming from each end of the sample is used.

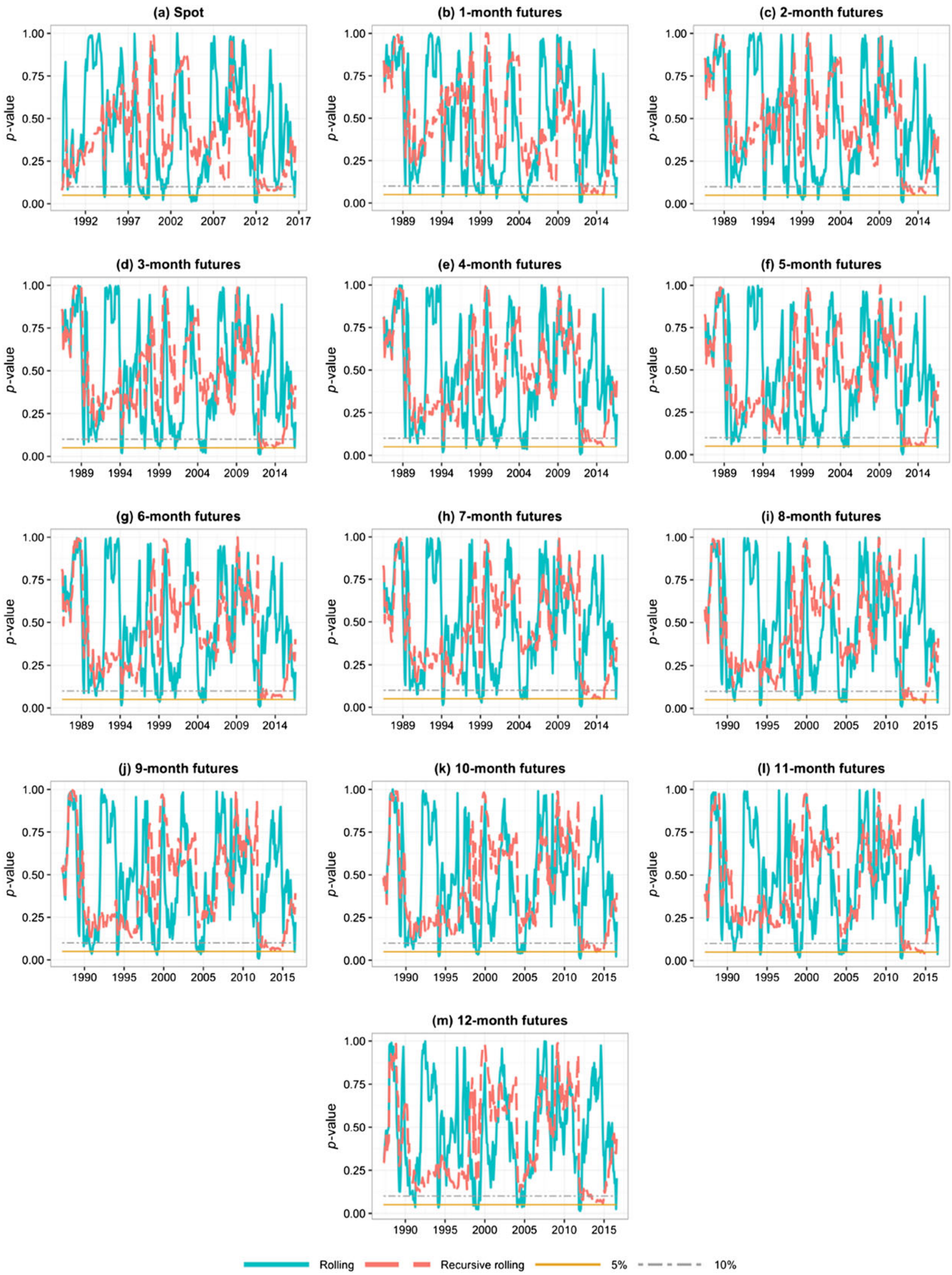
Although all three tests proposed by Andrews and Ploberger (1994) test the same null hypothesis, they differ in their choice of alternative hypotheses. The choice of which test to apply depends on the purpose of the test (see Andrews, 1993 and Andrews & Ploberger, 1994 for details). The results of the *Sup-F*, *Mean-F*, and *Exp-F* are shown in Table 3. According to the results reported in Table 3, all tests reject the null hypothesis of parameter constancy at the 5% level (at 10% only in one case) for the oil price equation, gold price equation, and the VAR system. Considering all the above-mentioned factors, it may be stated that Granger causality tests based on the VAR model estimated for oil prices and gold prices are not reliable because the parameters in the VAR model do not stay constant over the sample period.

The parameter instability tests reported in Table 3 establish a strong case for considering time-varying Granger causality tests because the parameters of all estimated VAR models show significant instability. Figure 2 displays the time-varying Wald test statistical results for causal effects from oil price to gold price, whereas Figure 3 shows the heteroskedasticity-consistent time-varying Wald test statistics results for causal effects running from oil price to gold price.[6] In Figure 4, we report the time-varying Wald test statistics results for causal effects running from gold price to oil price. Last, the heteroskedasticity-consistent Wald test results for Granger causality running from gold price to oil price are reported in Figure 5. In Figures 2–5, we report the bootstrap *p* values of the rolling and recursive rolling Wald tests obtained from a VAR model with a varying lag order and a minimum window size of 40. For each subsample, we select the optimal lag orders with a maximum lag order of 12 using the Bayesian Information Criterion. The *p* values of the tests are obtained using 1,000 bootstrap replications. The bootstrap *p* values of the rolling and recursive rolling Wald tests reported in Figures 2 and 4 are obtained under the assumption of homoskedastic residuals, whereas the bootstrap *p* values given in Figures 3 and 5 assume heteroskedastic residuals. Because the residuals of the estimated VAR models show heteroscedasticity (see Table 1), there might be differences between the standard Wald test results (Figures 2 and 4) and the heteroskedasticity-consistent Wald test results (Figures 3 and 5).

We will assess the noncausality tests at the 5% significance level; however, 10% critical values can also be considered to exercise caution against low test power due to the sample size in each rolling and recursive rolling subsample estimate. According to the results shown in Figures 2–5, the *p* values change substantially over the sample, indicating sharp structural breaks. The evidence from Figure 2 for the 13 cases—namely, at the spot and futures contracts at 1- to 12-month maturities—shows that the null hypothesis in which oil prices do not have predictive power for gold prices cannot be rejected at the 5% significance level for most of the sample, but there are subperiods where *p* values, particularly those of the rolling tests, fall below 5% and more frequently below 10%. The periods where there is causality from the oil price to gold price mostly fall between the 1990–2005 and 2012–2015 subperiods. The causality link from oil to gold markets were particularly prevalent approximately in the 1990–1992, 1994, 1997–1998, 2001–2002, 2004–2005, and 2012–2015 subperiods. In Figure 2, the rolling statistics indicates causal links more frequently than the recursive rolling statistics. Similar to the evidence given in Figure 2, the results from the heteroskedastic versions of the tests reported in Figure 3 indicate that the null hypothesis in which oil prices do not Granger causes gold prices for all cases cannot be rejected at the 5%significance level for most of the sample. However, like the homoskedastic Wald tests, heteroskedastic Wald tests indicate strong causality from oil markets to gold markets during the 1990–1992, 1994, 1997–1998, and 2001–2002 subperiods. Compared with the homoskedastic versions of the rolling and recursive rolling Wald causality tests, the heteroskedastic versions are in better agreement about the rejection periods of the noncausality hypotheses. Because the homoskedastic version of the recursive rolling test differs significantly from its heteroskedastic version compared with the rolling tests, we conclude that recursive rolling tests are more sensitive to heteroskedasticity, particularly in terms of its lower success rate in detecting causality links.

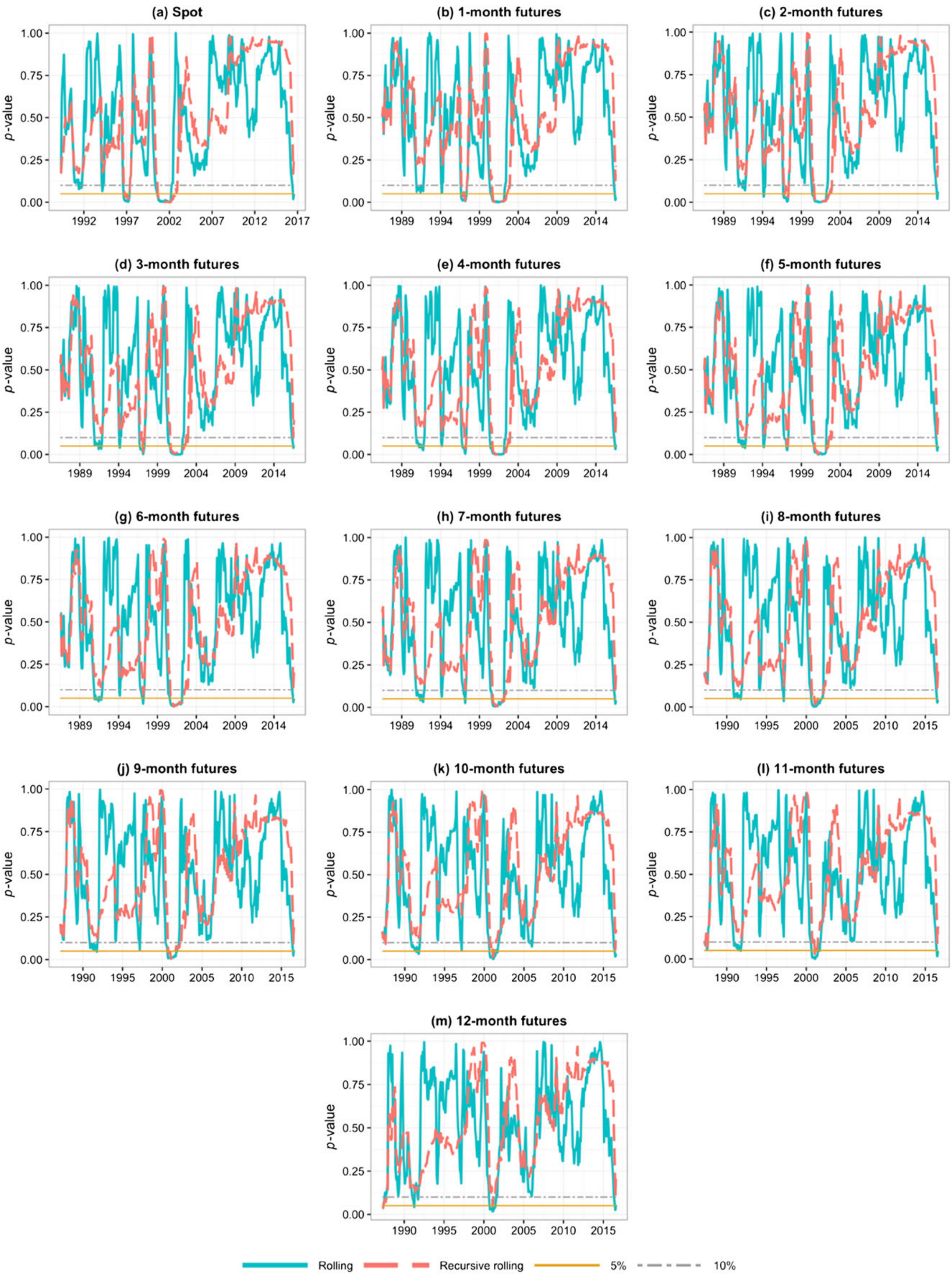
The bootstrap *p* values of the rolling and recursive rolling Wald tests statistics for the causal link from gold prices to oil prices in Figure 4 indicate that gold prices appear to have predictive power for oil prices only during the periods of 1997–1998 and 2007–2008. Parallel to the evidence in Figure 3, the bootstrap *p* values of the rolling and recursive rolling heteroskedasticity-consistent Wald tests statistics from Figure 5 do not reject the null hypothesis at the 5% significance level for almost all of the sample apart from two subperiods, where the null of Granger noncausality is rejected in the 1997–1998 and 2007–2008

[6]In this study, all rolling and recursive rolling Granger causality tests are performed using the using the bootstrap approach of Balcilar et al. (2010) to obtain the *p* values of the Wald statistics. Bootstrap approach is robust to small samples and integration–cointegration properties of the data (Mantalos, 2000; Mantalos & Shukur, 1998).



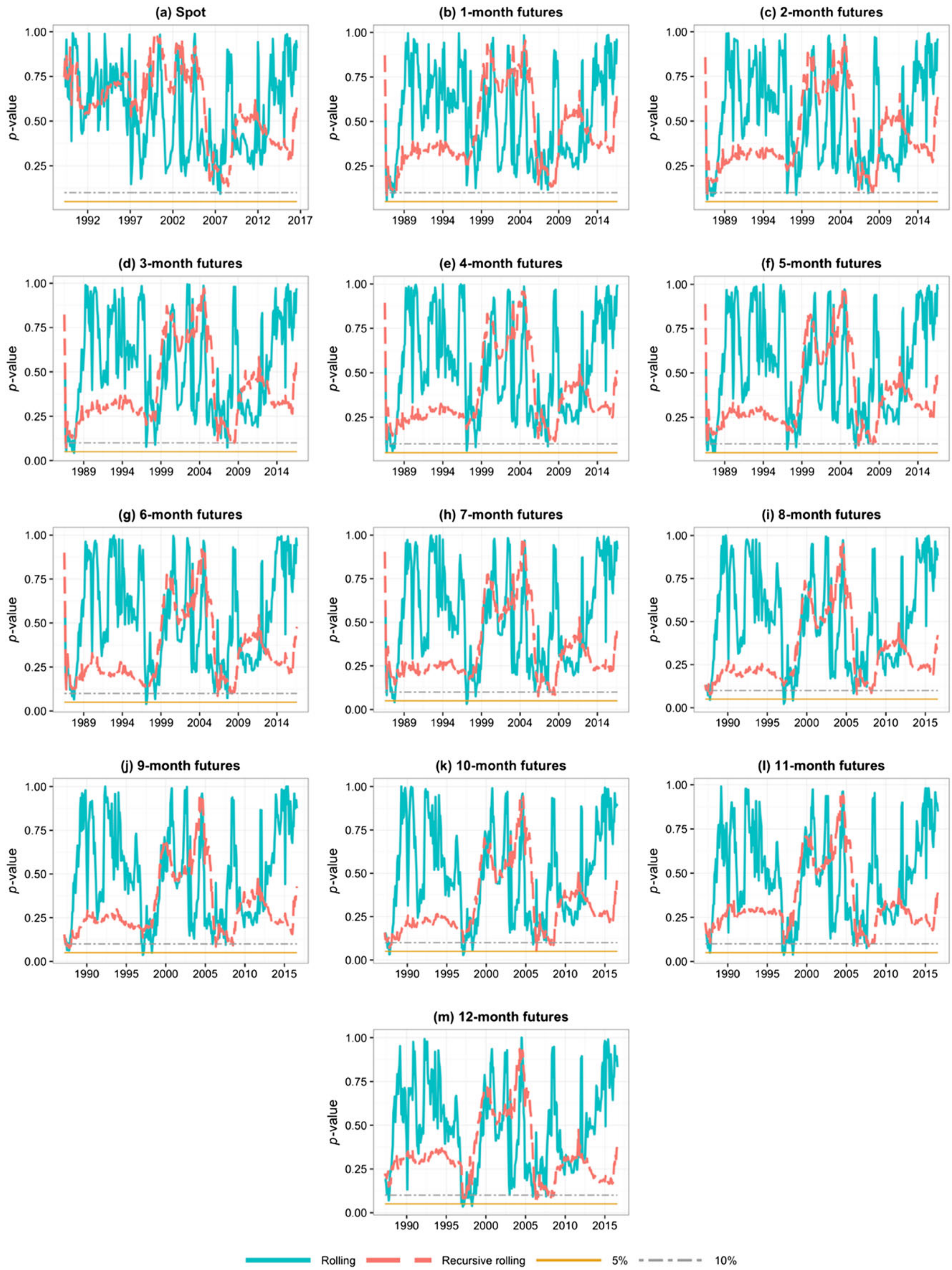
Note: The bootstrap p-values of the rolling and recursive rolling Wald tests are obtained from a VAR model with a varying lag order and a minimum window size of 40. For each sub-sample, the BIC is used to select the optimal lag orders with a maximum lag order of 12. The p-values of the tests are obtained using 1,000 bootstrap replications.

Figure 2. Wald tests for Granger causality running from the oil price to gold price



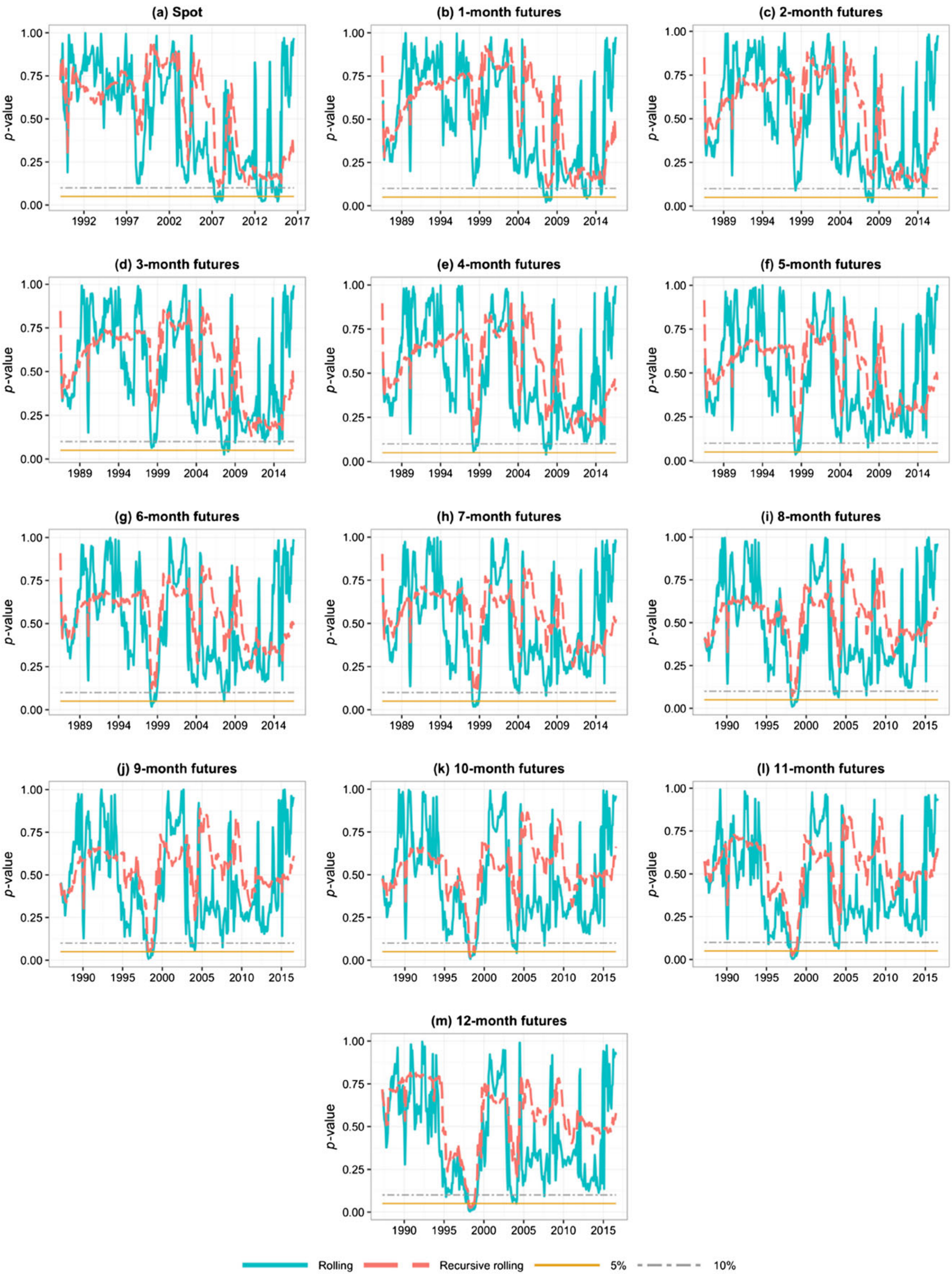
Note: See note to Figure 2.

Figure 3. Heteroskedasticity-consistent Wald tests for Granger causality running from the oil price to gold price



Note: See note to Figure 2.

Figure 4. Wald tests for Granger causality running from the gold price to oil price



Note: See note to Figure 2.

Figure 5. Heteroskedasticity-consistent Wald tests for Granger causality running from the gold price to oil price

subperiods. As for the Wald tests of Granger causality from oil to gold, heteroskedasticity-consistent rolling and recursive rolling tests are more consistent for the periods where they reject the null. Moreover, the recursive rolling approach appears more sensitive to heteroskedasticity than the rolling approach. The periods where there are causality links from gold price to oil price are around the late 1990s for 3- to 12-month maturity futures markets and the U.S. sub-prime mortgage crisis period from 2007 to 2008 for the spot and 1- to 3-month maturity futures markets. In addition to these subperiods, heteroskedasticity-consistent rolling tests in Figure 5 indicate two subperiods during which gold prices appear to have predictive power for oil prices for spot and 1- to 3-month maturity futures markets. These periods consist of the 2012–2013 and 2015 subperiods. Hurn et al. (2016) and Shi et al. (2016) state that the rolling method of Balcilar et al. (2010) and Balcilar and Ozdemir (2013a) is better at detecting structural breaks compared with the recursive rolling of Hurn et al. and Shi et al. The official National Bureau of Economic Research (NBER) recession period coincides with the U.S. sub-prime mortgage crisis period covering 2007–2008. Figures 2–5 show that the results from the rolling method of Balcilar et al. and Balcilar and Ozdemir detect structural breaks better than the recursive rolling of Hurn et al. and Shi et al.

In general, the heteroskedasticity-consistent version of the recursive rolling test appears to detect the same causality links as the rolling test, and the recursive rolling test shows greater sensitivity to heteroskedasticity. Because all our data show unconditional and conditional heteroskedasticity, it is more reasonable to consider the heteroskedasticity-consistent versions of tests to date stamp the periods of Granger causality. Both tests detect Granger causality from oil prices to gold prices in the 1990–1992, 1994, 1997–1998, and 2001–2002 subperiods. Each of these periods corresponds to significant oil price changes in crude oil due to influential events that had global impacts. In the 1990–1992 period, Iraq invaded Kuwait, leading to the first Persian Gulf War, which took place from August 1990 to February 1991. During this period, crude and other oil product prices rose significantly, exchange markets reacted wildly to any Middle East news events, and jet fuel prices increased to record spreads over other products due to increased defence demand. In 1994, oil prices likely increased due to three events. First, institutional investment funds in the United States shifted from equity and bond markets to cash and commodities. Second, oil production in Nigeria was disrupted because of the oil workers' strike. Third, led by Saudi Arabia, oil began to be traded in exchange for gold in the Middle East around 1994. The period 1997–1998 corresponds to East Asian crises that placed a strong downward pressure on oil prices. The crude oil price fell to $12 by the end of 1998, which was the lowest price since 1972. The East Asian rises were short lived, and strong growth due to new industrialization, particularly in China and India, caused a 38% rise in oil prices between November 1999 and November 2000. Oil prices continued to rise sharply until 2002. Analogously, the significant causality periods of 1997–1998 and 2007–2008 from gold to oil correspond to two large financial crises: the East Asian crisis in 1997–1998 and the sub-prime mortgage crisis in 2007–2008. The type of events correspond to the periods where causality links between the oil and gold markets indicate that the causality from oil to gold is related to large oil price changes, whereas the causality from gold to oil corresponds to financial crises where a huge amount of investment shifts took place between the gold and oil markets.

The empirical evidence for the relationship between gold and oil has been mixed and frequently disputed by economists until now. The first direct trade link between these commodities was established with selling oil in the Middle East in return for gold. In 1933, Saudi Arabia stated that it would sell oil only in return for gold, and this was an important turning point in terms of the relationship between gold and oil. The frequently repeated view concerning the gold–oil relationship is that there is a direct proportion between the prices of both commodities. Even though the rule does not hold under all conditions, the fact that there was a positive correlation between the Brent crude oil prices and the price of gold in the period between 1987 and 2012 can be regarded as an indicator confirming the generalization. Therefore, the tendency for the two prices to move jointly in recent years has aroused interest again to examine the links between the corresponding markets based on the assumption that gold and oil are major commodities and that the fluctuations in their prices have important implications for the real economy and financial markets. Huge increases in oil prices have been attributed to economic expansions, trade deficits, high inflation, high uncertainty in investments, and low stock and bond values. Gold, however, is considered a hedge against the risk of inflation and increasing financial market risks and is thus traded. The two commodities—oil and gold—can also influence the price of other commodities (see, e.g., Sari et al., 2010). The crude oil prices of West Texas Intermediate climbed to $135 in mid-July 2008, which was approximately $25 in the early 2000s, as seen in Figure 1. The increase in the price of gold also occurred until the first half of 2008. Such a joint movement was also observed in crude oil and gold prices during the financial crisis. Therefore, we empirically investigate the dynamic linkages between these markets. The evidence from the rolling method of Balcilar et al. (2010) and

Balcilar and Ozdemir (2013a) and the recursive rolling of Hurn et al. (2016) and Shi et al. (2016) used in this study shows that there is a bidirectional causal link between oil prices and gold prices for several subperiods for spot and futures prices at of 1- to 12-month maturities. Our findings provide valuable information on the causes of the oil–gold nexus by date stamping the periods of causality, showing that oil influences gold when there are large oil price changes and gold influences oil when there are large financial crises. Our results indicate that both the rolling and the recursive rolling methods show good performance in situations where there are multiple dual changes in the causal relationship between oil prices and gold prices over the sample period.

4. Conclusion

This study examined the causal nexus between oil prices and gold prices with the rolling and recursive rolling estimation approaches. The rolling and recursive rolling approaches are useful for investigating time-varying causal links. These approaches allow us to model parameter time variation to reflect changes in Granger causality without any assumptions of the change mechanism. The data series used in this article are monthly time series of oil prices and gold prices. What is novel in this study is the use of both a rolling Granger causality test and a recursive rolling Granger causality test. We first investigate the linear Granger causality between oil prices and gold prices. The findings from this test indicate that there is no predictive power between oil prices and gold prices. Next, we apply a battery of stability tests to the models from which the findings of full sample Granger causality tests are obtained. The results from the stability test show that the VAR models do not have stable parameters. Thus, Granger causality test results from the full sample VAR model estimated for the oil prices and gold prices series are not reliable. Taking into account this issue, we analyse the evaluation of the causal link between oil prices and gold prices over the study period using rolling and recursive rolling methods, the evidence from this paper shows two main conclusions. First, there is a bidirectional causal link between the series for all cases considered during several subperiods. Second, the rolling method proposed by Balcilar et al. (2010) and Balcilar and Ozdemir (2013a) is more robust in detecting the structural breaks in this context than the findings from recursive rolling method of Hurn et al. (2016) and Shi et al. (2016). Overall, the results from this study show that the dynamic relationships between oil price and gold price series will be sensitive to the frequency of the oil and gold series, time span, and method used. Therefore, the suggestion from this study is that it is very important that further studies be conducted on the causal relationships between oil prices and gold prices, as well as other cases, allowing policy changes and significant shifts in oil and gold and volatile periods caused by recessions and financial crises. Oil as commodity has been becoming more financialized over the last three decades. Like its role as a safe haven hedge instrument against the stock market risks, gold could be considered in hedging oil price risks. Our findings indicate that assuming noncausality between oil and gold might lead to dangers in using gold as an hedge instrument against the oil price risk.

References

Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica*, *61*, 821–856. https://doi.org/10.2307/2951764

Andrews, D. W. K., & Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica*, *62*, 1383–1414. https://doi.org/10.2307/2951753

Arora, V., & Shi, S. (2015). Energy consumption and economic growth in the United States. Working paper.

Balcilar, M., Bekiros, S., & Gupta, R. (2017a). The role of news-based uncertainty indices in predicting oil markets: A hybrid nonparametric quantile causality method. *Empirical Economics*, *53*(3), 879–889. https://doi.org/10.1007/s00181-016-1150-0

Balcilar, M., Gupta, R., & Pierdzioch, C. (2017b). On exchange-rate movements and gold-price fluctuations: Evidence for gold-producing countries from a nonparametric causality-in-quantiles test. *International Economics and Economic Policy*, *14*(4), 691–700. https://doi.org/10.1007/s10368-016-0357-z

Balcilar, M., & Ozdemir, Z. A. (2013a). The export-output growth nexus in Japan: A bootstrap rolling window approach. *Empirical Economics*, *44*(2), 639–660. https://doi.org/10.1007/s00181-012-0562-8

Balcilar, M., & Ozdemir, Z. A. (2013b). Asymmetric and time-varying causality between inflation and inflation uncertainty in G-7 countries. *Scottish Journal of Political Economy*, *60*, 1–41. https://doi.org/10.1111/sjpe.12000

Balcilar, M., & Ozdemir, Z. A. (2013c). The causal nexus between oil prices and equity market in the U.S.: A regime switching model. *Energy Economics*, *39*, 271–282. https://doi.org/10.1016/j.eneco.2013.04.014

Balcilar, M., Ozdemir, Z. A., & Arslanturk, Y. (2010). Economic growth and energy consumption causal nexus viewed through a bootstrap rolling window. *Energy Economics*, *32*(6), 1398–1410. https://doi.org/10.1016/j.eneco.2010.05.015

Bampinas, G., & Panagiotidis, T. (2015). Are gold and silver a hedge against inflation? A two century perspective. *International*

Review of Financial Analysis, *41*, 267–276. https://doi.org/10.1016/j.irfa.2015.02.007

Bildirici, M. E., & Turkmen, C. (2015). Nonlinear causality between oil and precious metals. *Resources Policy*, *46*, 202–211. https://doi.org/10.1016/j.resourpol.2015.09.002

Bilgin, M. H., Gogolin, F., Lau, M. C. K., & Vigne, S. A. (2018). Time-variation in the relationship between white precious metals and inflation: A cross-country analysis. *Journal of International Financial Markets, Institutions and Money.*, *56*, 55–70. https://doi.org/10.1016/j.intfin.2018.03.001

Das, D., Kumar, S. B., Tiwari, A. K., Shahbaz, M., & Hasim, H. M. (2018). On the relationship of gold, crude oil, stocks with financial stress: A causality-in-quantiles approach. *Finance Research Letters.*. https://doi.org/10.1016/j.frl.2018.02.030

Deutsch, M., Granger, C. W., & Teräsvirta, T. (1994). The combination of forecasts using changing weights. *International Journal of Forecasting*, *10*(1), 47–57.

Dolado, J. J., & Lütkepohl, H. (1996). Making Wald tests work for cointegrated VAR systems. *Econometric Reviews*, *15*, 369–386. https://doi.org/10.1080/07474939608800362

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics*, *7*, 1–26. https://doi.org/10.1214/aos/1176344552

Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, *55*(2), 251–276. https://doi.org/10.2307/1913236

Ghosh, S. (2011). Examining crude oil price—Exchange rate nexus for India during the period of extreme oil price volatility. *Applied Energy*, *88*(5), 1886–1889. https://doi.org/10.1016/j.apenergy.2010.10.043

Gil-Alana, L. A., Yaya, O. S., & Awe, O. O. (2017). Time series analysis of co-movements in the prices of gold and oil: Fractional cointegration approach. *Resources Policy*, *53*, 117–124. https://doi.org/10.1016/j.resourpol.2017.06.006

Granger, C. W. J. (1996). Can we improve the perceived quality of economic forecasts? *Journal of Applied Econometrics*, *11*, 455–473. https://doi.org/10.1002/(SICI)1099-1255(199609)11:5<455::AID-JAE408>3.0.CO;2-E

Granger, C. W. J., & Teräsvirta, T. (1993). *Modelling nonlinear economic relationships*. New York: Oxford University Press, ISBN 0-19-877320x.

Guilkey, D. K., & Salemi, M. K. (1982). Small sample properties of three tests for Granger- causal ordering in a bivariate stochastic system. *Review of Economics and Statistics*, *64*, 668–680. https://doi.org/10.2307/1923951

Hacker, R. S., & Hatemi-J, A. (2006). Tests for causality between integrated variables based on asymptotic and bootstrap distributions: Theory and application. *Applied Economics*, *38*, 1489–1500. https://doi.org/10.1080/00036840500405763

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, *57*, 357–384. https://doi.org/10.2307/1912559

Hammoudeh, S., & Yuan, Y. (2008). Metal volatility in presence of oil and interest rate shocks. *Energy Economics*, *30*, 606–620. https://doi.org/10.1016/j.eneco.2007.09.004

Hansen, B. E. (1992). The likelihood ratio test under nonstandard conditions: Testing the Markov switching model of GNP. *Journal of applied Econometrics*, *7*, 61–82. https://doi.org/10.1002/jae.3950070506

Hansen, B. E. (2001). The new econometrics of structural change: Dating breaks in U.S. labor productivity. *The Journal of Economic Perspectives*, *15*, 117–128. https://doi.org/10.1257/jep.15.4.117

Horowitz, J. L. (1994). Bootstrap-based critical values for the information matrix test. *Journal of Econometrics*, *61*, 395–411. https://doi.org/10.1016/0304-4076(94)90092-2

Hurn, S., Phillips, P. C. B., & Shi, S. (2016). Change detection and the causal impact of the yield curve. Cowles Foundation Discussion Paper No. 2059.

Jain, A., & Biswal, P. C. (2016). Dynamic linkages among oil price, gold price, exchange rate, and stock market in India. *Resources Policy*, *49*, 179–185. https://doi.org/10.1016/j.resourpol.2016.06.001

Jain, A., & Ghosh, S. (2013). Dynamics of global oil prices, exchange rate and precious metal prices in India. *Resources Policy*, *38*(1), 88–93. https://doi.org/10.1016/j.resourpol.2012.10.001

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, *12*, 231–254. https://doi.org/10.1016/0165-1889(88)90041-3

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica*, *59*, 1551–1580. https://doi.org/10.2307/2938278

Junttila, J., Pesonen, J., & Raatikainen, J. (2018). Commodity market based hedging against stock market risk in times of financial crisis: The case of crude oil and gold. *Journal of International Financial Markets, Institutions and Money*, *56*, 255–280.

Kanjilal, K., & Ghosh, S. (2017). Dynamics of crude oil and gold price post 2008 global financial crisis—New evidence from threshold vector error-correction model. *Resources Policy*, *52*, 358–365. https://doi.org/10.1016/j.resourpol.2017.04.001

Kim, M. H., & Dilts, D. A. (2011). The relationship of the value of the dollar, and the prices of gold and oil: A tale of asset risk. *Economics Bulletin*, *31*(2), 1151–1162.

Krolzig, H. M. (1999). *Statistical analysis of cointegrated VAR processes with Markovian regime shifts*. Manor Road, Oxford OX1 3UL, UK: University of Oxford.

Kumar, S. (2017). On the nonlinear relation between crude oil and gold. *Resources Policy*, *51*, 219–224. https://doi.org/10.1016/j.resourpol.2017.01.003

Le, T., & Chang, Y. (2011a). Oil and gold: Correlation or causation? Online at https://mpra.ub.uni-muenchen.de/31795/1/MPRA_paper_31795.pdf.

Le, T., & Chang, Y. (2011b). Dynamic relationships between the prices of oil, gold and financial variables in Japan: Abounds testing approach. Online at http://mpra.ub.uni-muenchen.de/33030/MPRA Paper No. 33030.

Lee, Y., Huang, Y., & Yang, H. (2012). The asymmetric long-run relationship between crude oil and gold futures. *Global Journal of Business Research*, *6*(1), 9–15.

Liao, S. J., & Chen, J. T. (2008). The relationship among oil prices, gold prices and individual industrial sub-indices in Thaiwan.

Working paper presented at international conference on business and information. Seoul, South Korea.

Mantalos, P. (2000). A graphical investigation of the size and power of the Granger-causality tests in integrated-cointegrated VAR systems. *Studies in Non-Linear Dynamics and Econometrics*, *4*, 17–33. https://doi.org/10.1162/108118200569162

Mantalos, P., & Shukur, G. (1998). Size and power of the error correction model cointegration test. A bootstrap approach. *Oxford Bulletin of Economics and Statistics*, *60*, 249–255. https://doi.org/10.1111/1468-0084.00097

Mei-Se, C., Shu-Jung, C. L., & Chien-Chiang, L. (2018). Time-varying co-movement of the prices of three metals and oil: Evidence from recursive cointegration. *Resources Policy*, *57*, 186–195. https://doi.org/10.1016/j.resourpol.2018.03.003

Melvin, M., & Sultan, J. (1990). South African political unrest, oil prices, and the time varying risk premium in the fold futures market. *Journal of Futures Markets*, *10*, 103–111. https://doi.org/10.1002/fut.3990100202

Mosconi, R., & Giannini, C. (1992). Non-causality in cointegrated systems: Representation, estimation, and testing. *Oxford Bulletin of Economics and Statistics*, *54*, 399–417. https://doi.org/10.1111/j.1468-0084.1992.tb00009.x

Narayan, P., Narayan, S., & Zheng, X. (2010). Gold and oil futures markets: Are markets efficient? *Applied Energy*, *87*, 3299–3303. https://doi.org/10.1016/j.apenergy.2010.03.020

Park, J. P., & Phillips, P. C. B. (1989). Statistical inference in regression with integrated process: Part 2. *Econometric Theory*, *5*, 95–131. https://doi.org/10.1017/S0266466600012287

Perron, P. (2006). Dealing with structural breaks. *Palgrave Handbook of Econometrics*, *1*, 278–352.

Phillips, P. C. B. (1995). Fully modified least squares and vector autoregression. *Econometrica*, *63*, 1023–1078. https://doi.org/10.2307/2171721

Psaradakis, Z., Ravn, M. O., & Sola, M. (2005). Markov switching causality and the money output relationship. *Journal of Applied Econometrics*, *20*, 665–683. https://doi.org/10.1002/jae.819

Sari, R., Hammoudeh, S., & Ewing, B. T. (2007). Dynamic relationships between oil and metal commodity futures prices. *Geopolitics of Energy*, *29*, 2–13.

Sari, R., Hammoudeh, S., & Soytas, U. (2010). Dynamics of oil price, precious metal prices, and exchange rate. *Energy Economics*, *32*, 351–362. https://doi.org/10.1016/j.eneco.2009.08.010

Sephton, P., & Mann, J. (2018). Gold and crude oil prices after the great moderation. *Energy Economics*, *71*, 273–281. https://doi.org/10.1016/j.eneco.2018.02.022

Shahbaz, M., Balcilar, M., & Ozdemir, Z. A. (2017). Does oil predict gold? A nonparametric causality-in-quantiles approach. *Resources Policy*, *52*, 257–265. https://doi.org/10.1016/j.resourpol.2017.03.004

Shi, S., Hurn, S., & Phillips, P. C. B. (2016). Causal change detection in possibly integrated systems: Revisiting the money-income relationship. Cowles Foundation Discussion Paper No. 2058.

Šimáková, J. (2011). Analysis of the relationship between oil and gold price. *Journal of Finance*, *51*(1), 651–662.

Sims, C. A., Stock, J. H., & Watson, M. W. (1990). Inference in linear time series models with some unit roots. *Econometrica*, *58*, 113–144. https://doi.org/10.2307/2938337

Soytas, U., Sari, R., Hammoudeh, S., & Hacıhasanoglu, E. (2009). World oil prices, precious metal prices and macroeconomy in Turkey. *Energy Policy*, *37*(12), 5557–5566. https://doi.org/10.1016/j.enpol.2009.08.020

Stern, D. I. (2000). A multivariate cointegration analysis of the role of energy in the US macroeconomy. *Energy Economics*, *22*(2), 267–283. https://doi.org/10.1016/S0140-9883(99)00028-6

Stock, J. H., & Watson, M. W. (1989). Interpreting the evidence on money-income causality. *Journal of Econometrics*, *40*, 161–181. https://doi.org/10.1016/0304-4076(89)90035-3

Swanson, N. R. (1998). Money and output viewed through a rolling window. *Journal of Monetary Economics*, *41*, 455–474. https://doi.org/10.1016/S0304-3932(98)00005-1

Teräsvirta, T. (1998). Modelling economic relationships with smooth transition regressions. In A. Ullah, & D. E. Giles (Eds.), *Handbook of applied economic statistics* (pp. 507–552). New York: Dekker.

Thoma, M. A. (1994). Subsample instability and asymmetries in money-income causality. *Journal of Econometrics*, *64*, 279–306. https://doi.org/10.1016/0304-4076(94)90066-3

Tiwari, A. K., & Sahadudheen, I. (2015). Understanding the nexus between oil and gold. *Resources Policy*, *46*, 85–91. https://doi.org/10.1016/j.resourpol.2015.09.003

Toda, H. Y., & Phillips, P. C. B. (1993). Vector autoregressions and causality. *Econometrica*, *61*, 1367–1393. https://doi.org/10.2307/2951647

Toda, H. Y., & Phillips, P. C. B. (1994). Vector autoregression and causality: A theoretical overview and simulation study. *Econometric Reviews*, *13*, 259–285. https://doi.org/10.1080/07474939408800286

Toda, H. Y., & Yamamoto, T. (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics*, *66*, 225–250. https://doi.org/10.1016/0304-4076(94)01616-8

Yamada, H., & Toda, H. Y. (1998). Inference in possibly integrated vector autoregressive models: Some finite sample evidence. *Journal of Econometrics*, *86*, 55–95. https://doi.org/10.1016/S0304-4076(97)00109-7

Zeileis, A., Leisch, F., Hornik, K., & Kleiber, C. (2005). Monitoring structural change in dynamic econometric models. *Journal of Applied Econometrics*, *20*, 99–121. https://doi.org/10.1002/jae.776

Zhang, Y. J., & Wei, Y. M. (2010). The crude oil market and the gold market: Evidence for cointegration, causality and price discovery. *Resources Policy*, *35*(3), 168–177. https://doi.org/10.1016/j.resourpol.2010.05.003