THE ROLE OF NEWS-BASED UNCERTAINTY INDICES IN PREDICTING OIL MARKETS: A HYBRID NONPARAMETRIC QUANTILE CAUSALITY METHOD

MEHMET BALCILAR [a, c, d*], STELIOS BEKIROS [a, b †], RANGAN GUPTA [a, d **]

[a] *IPAG Business School*, [b] *European University Institute (EUI)*, [c] *Eastern Mediterranean University*
[d] *University of Pretoria*

ABSTRACT

We emphasize the role of news-based economic policy and equity market uncertainty indices as robust drivers of oil price fluctuations. In that, we utilize a new hybrid nonparametric quantile causality methodology in order to investigate whether EPU and EMU uncertainty measures incorporate critical predictability for oil market returns and volatility. Based on an updated daily database spanning January 1986 to December 2014, we find that both measures present strong predictability over the entire distribution of oil around the median, yet more importantly for volatility forecastability covers the entire distribution except minor divergences in the tails. Therefore, an inherent heterogeneity is observed and an asymmetric pattern over the distribution of oil returns and its volatility exists with respect to uncertainty predictability.

JEL Codes: C32; C53; Q41
Keywords: Uncertainty; Oil markets; Volatility; Quantile causality

---

[* c] Department of Economics, Eastern Mediterranean University, Famagusta, Northern Cyprus, via Mersin 10, Turkey; E-mail address: mehmet@mbalcilar.net.

[†] ***Corresponding author***: [a] IPAG Business School, 184 Boulevard Saint-Germain, 75006 Paris, France.; Tel.: +33 01 53 63 36 00 ; Fax: +33 01 45 44 40 46 ; [b] Department of Economics, Via della Piazzuola; 43, I-50133, Florence, Italy; Tel.: +39 055 4685 916; Fax: +39 055 4685 902; E-mail address: stelios.bekiros@eui.eu.

[**d] Department of Economics, University of Pretoria, Pretoria, 0002, South Africa. E-mail address: rangan.gupta@up.ac.za.

1. INTRODUCTION

A recent strand of literature by Kang and Ratti (2013a, b) and Antonakakis *et al.* (2014) revisits the interrelationship of oil-price shocks with recessions and inflationary episodes in the US economy, following the seminal work of Hamilton (1983). Therefore, it is of paramount importance to determine the principal drivers of the oil market and develop efficient forecasting models for oil prices. Only until lately Bloom (2009), Colombo (2013) and Jones and Olson (2013) emphasized the role of economic policy uncertainty on real activity, which in turn affects oil-price fluctuations. Moreover, equity-market uncertainty feeds into oil-price movements as firm-based uncertainty related to hiring and investment affects decision-making about firm production efficiency. In addition, Kang *et al.* (2015) reported empirical evidence revealing a strong dependence between oil prices and stock market volatility. Under this framework, the objective of our paper is to investigate whether two novel news-based measures of economic policy uncertainty (EPU) and equity market uncertainty (EMU) developed by Baker *et al.* (2013) can predict oil returns and volatility. At the same time we consider the possibility that the oil market is also likely to drive those uncertainties in a reverse fashion, hence we employ a modified bivariate quantile causality test that builds upon the mixing conditions of causality-in-quantile as in Jeong *et al*. (2012) and of the higher-moment *k*-order nonparametric causality as in Nishiyama *et al*. (2011). For our purposes we utilize daily data of oil returns and of the EPU and EMU indices spanning the period January 2, 1986 to December 8, 2014.

While Kang and Ratti (2013a, b) and Antonakakis *et al*. (2014) report a weak and negative conditional mean-based evidence of EPU affecting monthly oil prices derived from structural vector autoregressive modelling, to the best of our knowledge our paper is the first attempt to analyse the importance of both uncertainty measures in forecasting oil returns and volatility over their entire conditional distribution. The causality-in-quantile approach employed in our study presents with the following novelties: firstly, it is robust to misspecification errors as it detects the underlying dependence structure between the examined time series; this could prove to be particularly important, as it is well known that high-frequency data display nonlinear dynamics. Secondly, via our methodology we test for causality that may exist in the tails of the joint distribution of the variables, thus not only for causality-in-mean (1st moment). Also we investigate causality-in-variance thereby volatility spillovers, as some times when causality in the conditional mean may not exist, yet higher order interdependencies emerge. The paper is organized as follows: section 2 describes the mathematical context of quantile and higher-moment nonparametric causality, whilst section 3 presents the empirical results. Finally, Section 4 concludes.

2. NONPARAMETRIC QUANTILE CAUSALITY TESTING

We present thereafter a novel methodology for the detection on nonlinear causality via a hybrid approach based on the Nishiyama *et al.* (2011) and Jeong *et al.* (2012) framework. We denote oil returns as (y_t) and EPU or EMU as (x_t). The quantile-based causality is defined as follows:

x_t does not cause y_t in the θ-quantile with respect to the lag-vector of $\{y_{t-1},...,y_{t-p},x_{t-1},...,x_{t-p}\}$ if

$$Q_\theta\{y_t \mid y_{t-1},...,y_{t-p},x_{t-1},...,x_{t-p}\} = Q_\theta\{y_t \mid y_{t-1},...,y_{t-p}\} \quad (1)$$

x_t is a prima facie cause of y_t in the θ^{th} quantile with respect to $\{y_{t-1},...,y_{t-p},x_{t-1},...,x_{t-p}\}$ if

$$Q_\theta\{y_t \mid y_{t-1},...,y_{t-p},x_{t-1},...,x_{t-p}\} \neq Q_\theta\{y_t \mid y_{t-1},...,y_{t-p}\} \quad (2)$$

where $Q_\theta\{y_t \mid \cdot\}$ is the θ^{th} quantile of y_t depending on *t* and $0 < \theta < 1$.

Let $Y_t \equiv (y_{t-1},...,y_{t-p})$, $Z_{t-1} \equiv (y_{t-1},...,y_{t-p},x_{t-1},...,x_{t-p})$, $V_t = (X_t, Z_t)$ and $F_{y_t|Z_{t-1}}(y_t, Z_{t-1})$ and $F_{y_t|Y_{t-1}}(y_t, Y_{t-1})$ denote the conditional distribution functions of y_t given Y_{t-1} and Z_{t-1} respectively. The conditional distribution $F_{y_t|Z_{t-1}}(y_t, Z_{t-1})$ is assumed to be absolutely continuous in y_t for almost all V_{t-1}. If we denote $Q_\theta(Z_{t-1}) \equiv Q_\theta(y_t \mid Z_{t-1})$ and $Q_\theta(Y_{t-1}) \equiv Q_\theta(y_t \mid Y_{t-1})$, we have $F_{y_t|Z_{t-1}}\{Q_\theta(Z_{t-1}) \mid Z_{t-1}\} = \theta$ with probability one. Consequently, the hypotheses to be tested based on definitions (1) and (2) are:

$$H_0 = P\{F_{y_t|Z_{t-1}}\{Q_\theta(Y_{t-1}) \mid Z_{t-1}\} = \theta\} = 1 \quad (3)$$

$$H_1 = P\{F_{y_t|Z_{t-1}}\{Q_\theta(Y_{t-1}) \mid Z_{t-1}\} = \theta\} < 1 \quad (4)$$

Jeong *et al.* (2012) employs the distance measure $J = \{\varepsilon_t E(\varepsilon_t \mid Z_{t-1}) f_z(Z_{t-1})\}$ where ε_t is the regression error term and $f_z(Z_{t-1})$ is the marginal density function of Z_{t-1}. The regression error ε_t emerges based on the null in (3), which can only be true if and only if

$E[1\{y_t \le Q_\theta(Y_{t-1}) \mid Z_{t-1}\}] = \theta$ or equivalently $1\{y_t \le Q_\theta(Y_{t-1})\} = \theta + \varepsilon_t$, where $1\{\cdot\}$ is an indicator function. Jeong *et al.* (2012) specify the distance function as follows:

$$J = E[\{F_{y_t|Z_{t-1}}\{Q_\theta(Y_{t-1}) \mid Z_{t-1}\} - \theta\}^2 f_Z(Z_{t-1})] \qquad (5)$$

In Eq. (3), it is important to note that $J \ge 0$ i.e., the equality holds if and only if H_0 in (5) is true, while $J > 0$ holds under the alternative H_1 in Eq. (4). Jeong *et al.* (2012) show that the feasible kernel-based test statistic for J has the following form:

$$\hat{J}_T = \frac{1}{T(1-1)h^{2p}} \sum_{t=1}^{T} \sum_{s \ne t}^{T} K\left(\frac{Z_{t-1} - Z_s}{h}\right) \hat{\varepsilon}_t \hat{\varepsilon}_s \qquad (6)$$

where $K(\cdot)$ is the kernel function with bandwidth h while $\hat{\varepsilon}_t$ is the estimate of the unknown regression error, which is estimated as follows:

$$\hat{\varepsilon}_t = 1\{y_t \le Q_\theta(Y_{t-1}) - \theta\} \qquad (7)$$

$\hat{Q}_\theta(Y_{t-1})$ is an estimate of the θth conditional quantile of y_t given Y_{t-1}. Below, we estimate $\hat{Q}_\theta(Y_{t-1})$ using the nonparametric kernel method as:

$$\hat{Q}_\theta(Y_{t-1}) = \hat{F}^{-1}_{y_t|Y_{t-1}}(\theta \mid Y_{t-1}) \qquad (8)$$

where $\hat{F}_{y_t|Y_{t-1}}(y_t \mid Y_{t-1})$ is the *Nadarya-Watson* kernel estimator given by:

$$\hat{F}_{y_t|Y_{t-1}}(y_t \mid Y_{t-1}) = \frac{\sum_{s \ne t} L(Y_{t-1} - Y_s) 1(Y_s \le Y_{t-1})}{\sum_{s \ne t} L \frac{(Y_{t-1} - Y_s)}{h}} \qquad (9)$$

with $L(\cdot)$ denoting the kernel function and h the bandwidth.

In an extension of the Jeong *et al.* (2012) framework, we develop a test for the 2nd moment. In particular, we want to test the volatility causality between EPU or EMU and oil returns. Causality in the mth moment implies causality in the kth moment for $k < m$. Firstly, we employ the nonparametric Granger quantile causality approach by Nishiyama *et al.* (2011). For a (y_t) process they assume that:

$$y_t = g(Y_{t-1}) + \sigma(X_{t-1})\varepsilon_t \qquad (10)$$

where $X_{t-1} = (x_{t-1}, x_{t-2}, ..., x_{t-p})$, ε_t is a white noise process and $g(\cdot)$, $\sigma(\cdot)$ are unknown functions that satisfy certain conditions for stationarity. However, this specification does not allow for Granger-type causality testing from x_t to y_t, but could possibly detect the "predictive power" from x_t to y_t^2 when $\sigma(\cdot)$ is a general nonlinear function. Hence, the Granger causality-in-variance definition does not require an explicit specification of squares for X_{t-1}. We re-formulate Eq. (10) into a null and alternative hypothesis for causality in variance as follows:

$$H_0 = P\{F_{y_t^2|Z_{t-1}}\{Q_\theta(Y_{t-1}) \mid Z_{t-1}\} = \theta\} = 1 \qquad (11)$$

$$H_1 = P\{F_{y_t^2|Z_{t-1}}\{Q_\theta(Y_{t-1}) \mid Z_{t-1}\} = \theta\} < 1 \qquad (12)$$

To obtain a feasible test statistic for testing the null in Eq. (10), we replace y_t in Eq. (6) - (9) with y_t^2. Incorporating the Jeong *et al.* (2012) approach we overcome the problem that causality in the conditional 1st moment (mean) may (or not) imply causality in the 2nd moment (variance) via the quantile measure now for higher than two order moments using the following model:

$$y_t = g(X_{t-1}, Y_{t-1}) + \varepsilon_t \qquad (13)$$

Thus, higher order quantile causality can be specified as:

$$H_0 = P\{F_{y_t^k|Z_{t-1}}\{Q_\theta(Y_{t-1}) \mid Z_{t-1}\} = \theta\} = 1 \quad \text{for } k = 1,2,...,K \qquad (14)$$

$$H_1 = P\{F_{y_t^k|Z_{t-1}}\{Q_\theta(Y_{t-1}) \mid Z_{t-1}\} = \theta\} < 1 \quad \text{for } k = 1,2,...,K \qquad (15)$$

Integrating the entire framework, we define that x_t *Granger causes* y_t *in quantile* θ *up to* K *th moment* utilizing Eq. (11) to construct the test statistic of Eq. (6) for each k. However, it can be shown that it is impossible to combine the different statistics for each $k = 1,2,...,K$ into one statistic for the joint null in Eq. (11) because the statistics are mutually correlated (Nishiyama *et al.*, 2011). To efficiently address this issue, we include a sequential-testing method as described Nishiyama *et al.* (2011) with some modifications. Firstly we test for the nonparametric Granger causality in the 1st moment $(k=1)$. Rejecting the null of non-causality means that we can stop and interpret this result as a strong indication of possible Granger quantile causality-in-variance. Nevertheless, failure to reject the null for $k=1$, does not automatically leads to no-causality in the 2nd moment, thus we can still construct the tests for $k=2$. Finally, we can test the existence of causality- in-variance, or the causality-in-mean and variance successively. The empirical implementation of causality testing via quantiles entails specifying three important choices: the bandwidth h, the lag order p, and the kernel type for $K(\cdot)$ and $L(\cdot)$ in Eq. (6) and (9) respectively. In our study, the lag order (9 and 5 respectively) is determined using the Schwarz Information Criterion (SIC) under a VAR comprising oil returns and EPU or EMU respectively. The bandwidth value is selected using the least squares cross-validation method. Lastly, for $K(\cdot)$ and $L(\cdot)$ we employ Gaussian-type kernels.

3. DATA ANALYSIS AND EMPIRICAL RESULTS

We empirically investigate the ability of EPU and EMU in predicting oil returns and their volatility over various quantiles, using data spanning the period 2nd January, 1986 to 8th December, 2014. The start and end dates of the sample are purely driven by data availability. We use the EMU and EPU indices developed by Baker *et al.* (2013) as two measures of uncertainty for the US economy. In particular, the daily news-based EPU index uses newspaper archives from the Access World News Bank service. The primary source comprises the number of articles containing at least one term from each of 3 sets of terms namely, "*economic / economy*", "*uncertain / uncertainty*" and "*legislation / deficit / regulation / congress / Federal Reserve / White House*"[1]. Using the same news source, the EMU index incorporates articles containing the aforementioned terms and one or more of the terms "*equity market / equity price / stock market*". We work with the natural logarithmic values of EPU and EMU. Additionally, we use the daily spot price of West Texas Intermediate (WTI) crude as a proxy for the oil market, derived the database of the Federal Reserve Bank of St. Louis.[2] We express the oil prices as returns i.e., the natural logarithmic difference expressed as percentages, to ensure stationarity. The total sample size provides with 7299 observations. Overall, we use the news-based measure of EMU index instead of

[1] Further details appear at: http://www.policyuncertainty.com/us_daily.html and http://www.policyuncertainty.com/equity_uncert.html.

[2] The FRED database is provided in http://research.stlouisfed.org/fred2/.

the VIX - a popular measure of the implied volatility of S&P 500 index options - in an attempt to surpass and outreach the basic index of uncertainty beyond the financial markets. In that, the EMU news articles-based measure is wider, whilst at the same time in terms of its relationship with VIX is directly comparable.[3]

The standard unit root tests reveal that while oil prices are unit root processes, the returns are stationary. Furthermore, EMU and EPU measures are found to be stationary as well. We used the Brock *et al.* (BDS, 1996) test on the residuals of AR(1) models of oil returns and EPU and EMU indices as well as of two VAR(1) models comprising oil returns and EPU or EMU. The null hypothesis of serial dependence is strongly rejected at 1% level of significance across various dimensions. Those results provide evidence of nonlinearities in the data. Finally, it is shown that oil returns are skewed to the left while EPU and EMU are rightly skewed, hence all variables following non-normal distributions. Moreover, we investigated whether EPU and EMU leads or lags oil returns by conducting the standard linear Granger causality testing based on a VAR model with nine (five) lags for EPU (EMU). The null hypothesis of EPU (EMU) not Granger causing oil returns was rejected at the 1% (10%) level of significance. Next we employed parameter (in)stability testing developed by Andrews (1993) and Andrews and Ploberger (1994) for the oil returns based on two VAR models. The null of stability was rejected at 1% level of significance by all test-stats i.e., *Sup-F*, *Exp-F* and *Ave-F*. The results were also corroborated by the Bai and Perron (2003) test of multiple structural breaks, which detected four and three breaks in the oil returns series for the EPU- and EMU-based VARs respectively[4]. Hence, in light of the strong evidence of structural instability, we cannot robustly rely on the linear Granger causality approach and the nonparametric quantile causality methodology presented above is utilized.

Figures 1a-b and 2a-b provide a graphical representation of the predictive ability of EPU and EMU for oil returns and volatility. The results are quite similar for EPU and EMU across the various quantiles, with the null hypothesis of no-causality being rejected for quantiles below 0.40 and above 0.55, and below 0.45 and above 0.55 respectively. As far as the volatility is concerned, the situation appears to be similar, with the exception of a slight difference at the higher quantiles of the distributions. Indeed, whilst for EPU the null is rejected for all quantiles above 0.15, for EMU the null is rejected between quantiles 0.20 and 0.85. Overall, the results indicate that EPU and EMU predictability vis-à-vis the oil returns is

[3] As indicated in http://www.policyuncertainty.com/equity_uncert.html, the EMU exhibits a contemporaneous daily correlation with the VIX of over 40%.
[4] Complete details of these results are available upon request from the authors.

high except around a region close to the median, whereas for volatility the predictability virtually extends throughout the entire distribution, with minor exceptions in the tails.[5]

4. CONCLUSIONS

There exists an important stream of literature relating oil-price movements (shocks) with recessions and inflationary episodes in the US economy. Taking this into account, it is of utmost importance to determine the variables that "drive" the oil markets. Recent works emphasize the role of uncertainty – derived from economic policy and equity markets - as proxies of oil price fluctuations. In light of new evidence, our objective was to analyse whether recently developed news-based measures, namely economic policy uncertainty (EPU) and equity market uncertainty (EMU) indices, can enhance the predictability of returns and volatility of oil prices. For this purpose, we introduced a novel methodology for the detection on nonlinear causality via a hybrid approach based on the Nishiyama *et al.* (2011) and Jeong *et al.* (2012) framework. Based on daily WTI oil returns and the EPU and EMU indices, we found that for the period January 1986 till December 2014 both measures present strong predictability over the entire distribution of oil returns around the median, yet more importantly for volatility the predictability covers the entire distribution except minor divergences in the tails. Consequently, uncertainty variables are likely to predict returns under turbulent oil markets, whilst volatility presents further forecastability in "normal" periods as well. It seems that an inherent heterogeneity is observed leading to an asymmetric pattern over the distribution of oil returns and its volatility with respect to uncertainty predictability.

[5] We obtained qualitatively similar results when we used oil returns and volatility based on the Brent crude price covering the period of 21st May, 1987 to 8th December, 2014. These results are available upon request from the authors.

REFERENCES

Andrews, Donald W. K. (1993). Tests for parameter instability and structural change with unknown change point, *Econometrica*, 61, 821–856.

Andrews, Donald W. K. and W. Ploberger (1994). Optimal tests when a nuisance parameter is present only under the alternative," *Econometrica*, 62, 1383–1414.

Antonakakis, N., Chatziantoniou, I. and Filis, G., (2014). Dynamic spillovers of oil price shocks and economic policy uncertainty. *Energy Economics* 44, 433–447.

Baker, S., Bloom, N., Davis, S., (2013). Measuring economic policy uncertainty. *Chicago Booth Research Paper No.13-02*.

Bai, Jushan and Pierre Perron (2003a). Computation and analysis of multiple structural change models, *Journal of Applied Econometrics*, 6, 72–78.

Bloom, N., (2009). The impact of uncertainty shocks. *Econometrica* 77, 623-685.

Brock, W., Dechert, D., Scheinkman, J. and LeBaron, B. (1996). A test for independence based on the correlation dimension. *Econometric Reviews*, 15, 197–235.

Colombo, V., (2013). Economic policy uncertainty in the US: Does it matter for the Euro area? *Economics Letters*, 121, 39-42.

Hamilton, J. D., (1983). Oil and the macroeconomy since World War II. *Journal of Political Economy*, 91, 228-248.

Jeong, K., Härdle, W.K., and Song, S. (2012). A consistent nonparametric test for causality in quantile. *Econometric Theory*, 28, 861 – 887.

Jones, P.M., and Olson, E., (2013). The time-varying correlation between uncertainty, output, and inflation: Evidence from a DCC-GARCH model. *Economics Letters*, 118, 33-37.

Kang, W. and Ratti, R. A. (2013a). Structural oil price shocks and policy uncertainty. *Economic Modelling* 35, 314-319.

Kang, W. and Ratti, R.A., (2013b). Oil shocks, policy uncertainty and stock market return. *International Financial Markets, Institutions and Money* 26, 305-318.

Kang, W., Ratti, R.A., Yoon, K.H., (2015). The impact of oil price shocks on the stock market return and volatility relationship. *Journal of International Financial Markets, Institutions and Money* 34, 41-54.

Nishiyama, Y., Hitomi, K., Kawasaki, Y., and Jeong, K. (2011). A consistent nonparametric test for nonlinear causality - Specification in time series regression. *Journal of Econometrics*, 165, 112-127.

FIGURE 1A: QUANTILE CAUSALITY RESULTS FOR THE H_0: *EPU DOES NOT GRANGER - CAUSE OIL RETURNS*

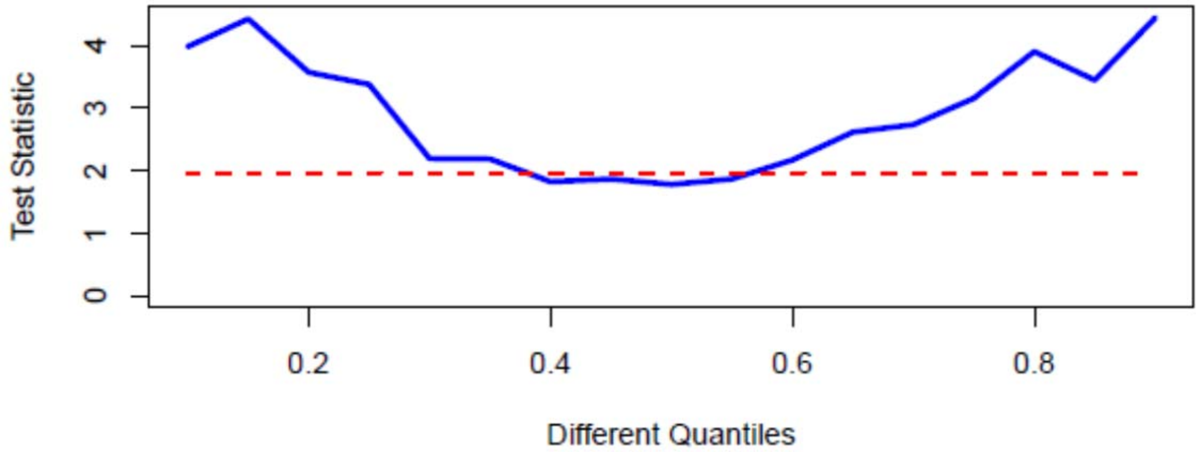

FIGURE 1B: QUANTILE CAUSALITY RESULTS FOR THE H_0: *EPU DOES NOT GRANGER - CAUSE OIL RETURN VOLATILITY*

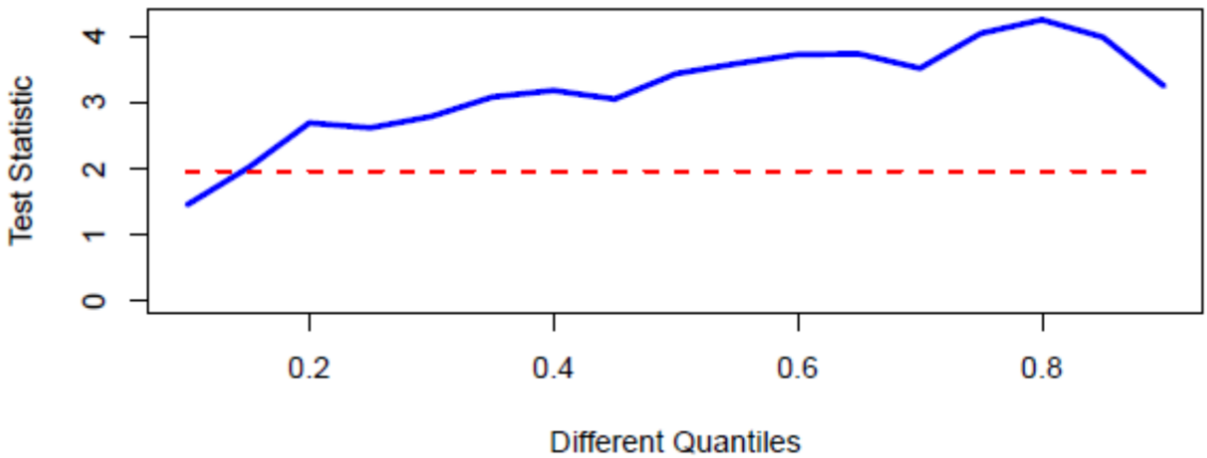

FIGURE 2A: QUANTILE CAUSALITY RESULTS FOR THE H_0: *EMU DOES NOT GRANGER - CAUSE OIL RETURNS*

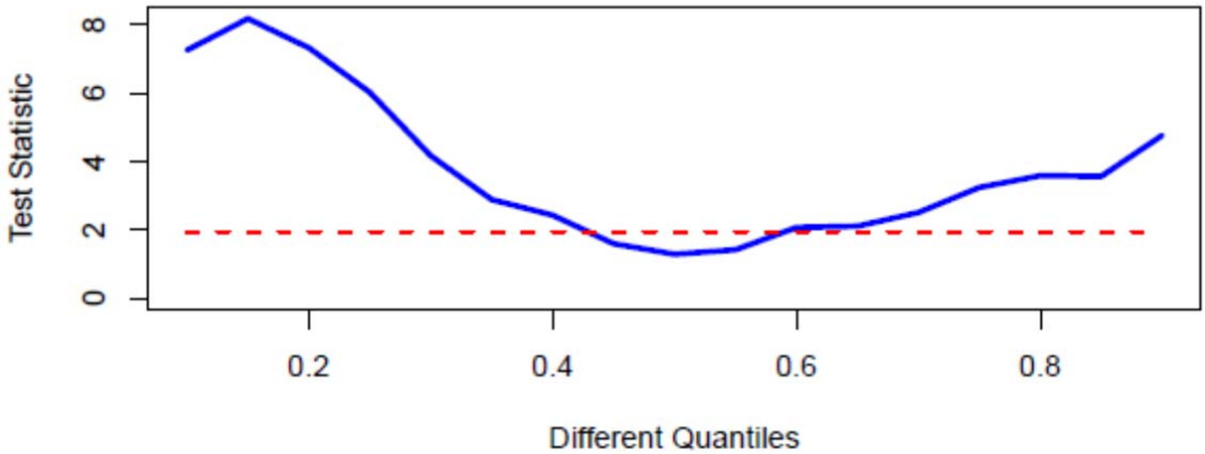

FIGURE 2B: QUANTILE CAUSALITY RESULTS FOR THE H_0: *EMU DOES NOT GRANGER - CAUSE OIL RETURN VOLATILITY*