The Dynamic Impact of Uncertainty in Causing and Forecasting the Distribution of Oil Returns and Risk

Giovanni Bonaccolto
University of Padova
Massimiliano Caporin
University of Padova
Rangan Gupta
University of Pretoria
Working Paper: 2015-64
September 2015
The Dynamic Impact of Uncertainty in Causing and Forecasting the Distribution of Oil Returns and Risk

G. Bonaccolto¹, M. Caporin², and R. Gupta³

¹Department of Statistical Sciences, University of Padova, via C. Battisti 241, 35121 Padova, Italy. 
Email: bonaccolto@stat.unipd.it

²Department of Economics and Management “Marco Fanno”, University of Padova, via del Santo 33, 35123 Padova, Italy. Email: massimiliano.caporin@unipd.it

³Department of Economics, University of Pretoria, 0002 Pretoria, South Africa. Email: Rangan.Gupta@up.ac.za

Abstract

The aim of the work is to analyse the relevance of recently developed news-based measures of economic policy uncertainty and equity market uncertainty in causing and predicting the conditional quantiles and distribution of the crude oil variations, defined both as returns and squared returns. For this purpose, on the one hand, we study the causality relations in quantiles through a non-parametric testing method; on the other hand, we forecast the conditional distribution on the basis of the quantile regression approach and the predictive accuracy is evaluated by means of several suitable tests. Given the presence of structural breaks over time, we implement a rolling window procedure to capture the dynamic relations among the variables.

Keywords: Granger Causality in Quantiles; Quantile Regression; Forecast of Oil Distribution; Forecast Evaluation.

JEL codes: C58, C32, C53, Q02, Q45.

1 Introduction

Following the seminal work of Hamilton (1983), a large literature exists that connects movements in oil returns and its volatility with recessions and inflationary episodes in the US economy (e.g., see Elder and Serletis (2010), Kang and Ratti (2013b,a), Antonakakis et al. (2014) for detailed reviews). Hamilton (2008) indicates that nine of ten recessions in the US since World War II have been preceded by an increase in oil prices (Hamilton, 2008). Interestingly, Hamilton (2009) even goes as far as arguing that a large proportion of the recent downturn in the US GDP during the “Great Recession” can also be attributed to the oil price shock of 2007-2008.
In turn, this implies that it is of paramount importance to determine the variables that drives the oil market to properly model and forecast, both returns and volatility of oil spot prices. In this regard, a recently growing literature emphasizes the role of economic policy uncertainty on real activity (e.g., see Bloom (2009), Colombo (2013), Jones and Olson (2013), Mumtaz and Zanetti (2013), Karnizova and Li (2014), Jurado et al. (2015) for detailed reviews), which, in turn, affects oil-price movements (Kang and Ratti, 2013b,a; Antonakakis et al., 2014; Aloui et al., 2015). Equity-market uncertainty also feeds into oil-price movements because, as Bloom (2009)’s firm-based theoretical framework notes, equity-market uncertainty affects hiring and investment and, hence, production decisions of firms. In this regard, empirical evidence relating oil price movements and stock market volatility can be found in Kang et al. (2015).

Against this backdrop, the objective of this paper is to analyse whether recently developed news-based measures of economic policy uncertainty (EPU) and equity market uncertainty (EMU) by Baker et al. (2013) can predict, both in- and out-of-sample, returns and volatility of oil. Realizing the possibility that the oil market is also likely to drive these uncertainties (see e.g. Kang and Ratti (2013b,a), Antonakakis et al. (2014)), we employ a modified bi-variate quantile causality-based model (for prediction and forecasting, as developed by Balcilar et al. (2015)), which combines the causality in quantile test of Jeong et al. (2012), with the $k$-th order nonparametric Granger causality test of Nishiyama et al. (2011) for our purpose, using daily data on oil returns, EPU and EMU, covering the period 02/01/1986-23/04/2015.

Conditional mean-based evidence of EPU (mildly and negatively), affecting oil price from structural vector autoregressive (SVAR) models, can be found in Kang and Ratti (2013b,a) and Antonakakis et al. (2014), and confirmed using copula models by Aloui et al. (2015). To the best of our knowledge, our paper is the first attempt to analyse the importance of both EPU and EMU in forecasting both in- and out-of-sample oil returns and its volatility over the entire conditional distribution of oil returns and volatility. The nonparametric causality in quantile test employed in our study for both in-sample and out-of-sample forecasting has the following novelties: first, the test is robust to functional misspecification errors and can detect general dependence between time series. This is particularly important in our application, since it is well known that high-frequency data display nonlinear dynamics. Second, the test statistic does not only test for causality in the mean, it also tests for causality that may exist in the tail area of the joint distribution of the series. Third, the test easily lends itself to test for causality in variance. Testing for causality in variance allows us to test for the volatility spillover phenomenon, 

---

1The Brock et al. (1996) test applied to the residuals recovered from autoregressive models fitted to oil returns and natural logarithms of EPU and EMU, as well as to the vector autoregressive models comprising of oil returns and logarithms of EPU or EMU, reject the null hypothesis of serial dependence at 1% level of significance across various dimensions. These results provide strong evidence of nonlinearity in the data. Complete details of these tests are available upon request from the authors.

2Our data showed that oil returns is skewed to the left while, EPU and EMU are skewed to the right, with all the three variables having non-normal distributions. Complete details on the summary statistics of the three variables are available upon request from the authors.
since, at times, causality in conditional mean (first moment) may not exist, but there may be second
or higher order causality. Moreover, the use of quantile-based methods, allows analysing the causality
structure depending on the volatility state (high versus low). Understandably, given the structure of
the model employed, it is easily tenable to forecasting the entire conditional distribution out-of-sample
for both oil returns and volatility, using a recursive or rolling estimation of the model. Besides, the
evaluation of asymmetric effects produced by EPU and EPU on the oil movements is important in
revealing the states where uncertainty assumes critical relevance.

At this stage, however, two related papers require mentioning. First, Bekiros et al. (2015) analysed
the importance of the EPU in forecasting oil returns over the in- and post crisis periods, using a wide
variety of constant parameter and time-varying parameter VAR models. The authors depict that EPU
matters in point forecasts of oil returns, but only when one allows for time-varying parameters (with
stochastic volatility in the error structure) in the VAR model. And second, Balcilar et al. (2015) who
develops the framework we use in this paper, used the model to analyse in-sample causality running
from EPU and EMU to oil returns and volatility. They concluded that, for oil returns, EPU and
EMU has strong predictive power over the entire distribution barring regions around the median,
but for volatility, the predictability virtually covers the entire distribution, with some exceptions in
the tails. Our contribution primarily involves extending the paper by Balcilar et al. (2015) to out-
of-sample density forecasting of oil returns and its volatility using a rolling window scheme. Note
that, in the process, we are also able to provide a time-varying approach to the in-sample quantile
causality for both oil returns and its volatility. This is important, given that we detect breaks in
their respective conditional distributions, and hence, full-sample quantile causality could be possibly
misleading. Further, unlike Bekiros et al. (2015), where the authors only concentrate on point forecast
of oil returns, we are able to analyse density forecast for both returns and volatility of oil returns. This
again is more informative than point forecasts, since we are able to understand the role of EPU and
EMU in forecasting oil returns and volatility at different phases (bearish, normal and bullish) of the
oil market. Hence, our contribution primarily involves looking at out of-sample density forecasts for
oil returns and its volatility using the information content of measures of policy and equity market
uncertainties at the highest possible (daily) frequency. The importance of our contribution can be
justified by the suggestion made by Campbell (2008): “The ultimate test of any predictive model is
its out-of-sample performance”. The rest of the paper is organized as follows: Section 2 presents the
details of the methodologies pursued, while in Section 3 we describe the data and the rolling window
procedure. Section 4 presents the results and Section 5 concludes, with an economic discussion of the
results obtained.
2 Causality and forecasting methods

2.1 Causality in quantiles

Let \( \{y_t\}_{t \in T} \) be the time series of the oil returns. Differently, the logarithm of the Economic Policy Uncertainty (EPU) and the logarithm of the Equity Market Uncertainty (EMU) are denoted, respectively, as \( \{x_{1,t}\}_{t \in T} \) and \( \{x_{2,t}\}_{t \in T} \).\(^3\) In the present section we describe a method which studies the causality relations in a bivariate framework. For simplicity of notation, in the following we use \( x_t \) in place of \( x_{1,t} \) \((x_{2,t}) \) when we study the causality implications of EPU (EMU) on \( y_t \). First, we make use of the following notations: \( Y_{t-1} = (y_{t-1}, \ldots, y_{p}) \), \( X_{t-1} = (x_{t-1}, \ldots, x_{q}) \), \( Z_{t-1} = (y_{t-1}, \ldots, y_{p}, x_{t-1}, \ldots, x_{q}) \), for \((p, q) > 1\); moreover, we denote by \( F_{y_t|Z_{t-1}}(y_t|Z_{t-1}) \) and \( F_{y_t|Y_{t-1}}(y_t|Y_{t-1}) \) the distributions of \( y_t \), conditional on \( Z_{t-1} \) and \( Y_{t-1} \), respectively. We assume the distribution of \( y_t \) is absolutely continuous in \( y \) for almost all \( \nu = (\mathcal{Y}, \mathcal{Z}) \). To simplify the notation, for any \( \tau \in (0, 1) \), we denote by \( Q_{\tau}(Z_{t-1}) \equiv Q_{\tau}(y_t|Z_{t-1}) \) and \( Q_{\tau}(Y_{t-1}) \equiv Q_{\tau}(y_t|Y_{t-1}) \) the \( \tau \)-th quantiles of \( y_t \) conditional to \( Z_{t-1} \) and \( Y_{t-1} \), respectively.

Granger (1989) defines the causality in mean (the well-known Granger causality) by means of a comparison between expected values computed conditioning to two different sets. We thus say that \( x_t \) does not cause \( y_t \) in mean with respect to \( Z_{t-1} \) if \( E[y_t|Z_{t-1}] = E[y_t|Y_{t-1}] \). Differently, \( x_t \) is a prima facie cause in mean of \( y_t \) with respect to \( Z_{t-1} \) if \( E[y_t|Z_{t-1}] \neq E[y_t|Y_{t-1}] \).

To define the Granger causality in quantiles, we follow Jeong et al. (2012) that defines the causality as follows: \( x_t \) does not cause \( y_t \) in its \( \tau \)-th quantile, with respect to \( Z_{t-1} \), if \( Q_{\tau}(Z_{t-1}) = Q_{\tau}(Y_{t-1}) \). On the other hand, \( x_t \) is a prima facie cause in the \( \tau \)-th quantile of \( y_t \) with respect to \( Z_{t-1} \), if \( Q_{\tau}(Z_{t-1}) \neq Q_{\tau}(Y_{t-1}) \).

The quantile causality definition leads to the identification of hypotheses we could test. Our interest lies in the detection of causality, and, similarly to the tests for Granger causality, we associate the null hypothesis to the absence of causality. As a result, the system hypotheses to be tested is:

\[
\begin{align*}
H_0 & : P[F_{y_t|Z_{t-1}}(Q_{\tau}(Y_{t-1})|Z_{t-1}) = \tau] = 1 \\
H_1 & : P[F_{y_t|Z_{t-1}}(Q_{\tau}(Y_{t-1})|Z_{t-1}) = \tau] < 1
\end{align*}
\]

(1)

In order to test the hypotheses given in (1), Jeong et al. (2012) suggest the use of the a specific distance measure:

\[
J_T = E \left[ |F_{y_t|Z_{t-1}}(Q_{\tau}(Y_{t-1})|Z_{t-1}) - \tau|^2 g_{Z_{t-1}}(Z_{t-1}) \right],
\]

(2)

where \( g_{Z_{t-1}}(Z_{t-1}) \) denotes the marginal density function of \( Z_{t-1} \). Notably, \( J_T \geq 0 \) with equality\(^3\)Detailed descriptions of the oil returns, EPU and EMU are given in Section 3.1.
holding if $H_0$ in (1) is true, while under the alternative $H_1$ we have a strict inequality.

Jeong et al. (2012) proposed the evaluation of the distance function by the following feasible kernel-based estimator:

$$
\hat{J}_T = \frac{1}{T(T-1)h^m} \sum_{t=1}^{T} \sum_{s \neq t} K \left( \frac{Z_{t-1} - Z_{s-1}}{h} \right) \tilde{\epsilon}_t \tilde{\epsilon}_s,
$$

where $m = p + q$, $K(\cdot)$ is the kernel function with bandwidth $h$, whereas $\tilde{\epsilon}_t$ is defined as

$$
\tilde{\epsilon}_t = 1_{\{y_t \leq \tilde{Q}_\tau(Y_{t-1})\}} - \tau,
$$

with $1_{\{\cdot\}}$ denoting the indicator function taking value 1 if the condition in $\{\cdot\}$ is true and zero otherwise.

Jeong et al. (2012), set $\tilde{Q}_\tau(Y_{t-1})$ equal to $\tilde{F}^{-1}_{y_t|Y_{t-1}}(\tau|Y_{t-1})$, where

$$
\tilde{F}_{y_t|Y_{t-1}}(y_t|Y_{t-1}) = \frac{\sum_{s \neq t} C_{t-1,s-1} 1_{\{y_s \leq y_t\}}}{\sum_{s \neq t} C_{t-1,s-1}}
$$

is the Nadaraya-Watson kernel estimator of $F_{y_t|Y_{t-1}}(y_t|Y_{t-1})$, with the kernel function $C_{t-1,s-1} = C(Y_{t-1} - Y_{s-1})/a$, and $a$ is the bandwidth.

Finally, Jeong et al. (2012) prove that, given $\sigma_\epsilon^2(Z_{t-1}) = \tau(1 - \tau)$ and a set of additional assumptions, $T h^{m/2} \hat{J}_T \overset{d}{\to} \mathcal{N}(0, \sigma_0^2)$, where

$$
\sigma_0^2 = 2E \left[ \sigma_\epsilon^4(Z_{t-1}) g_{Z_{t-1}}(Z_{t-1}) \right] \left( \int K^2(u) du \right).
$$

Therefore, the test for the presence of causality in a given quantile corresponds to a significance test of the quantity in (3) whose standard error depends on the sample estimator of the variance reported in equation (6).

More recently, Balcilar et al. (2015) extended the approach introduced by Jeong et al. (2012), and developed a test for the causality in quantiles but with a focus on the second moment of $y_t$. This novel approach allows testing for the presence of quantile causality when considering the density of the risk or of the dispersion characterising the variable $y_t$. We refer to this type of causality as quantile causality in variance.

Balcilar et al. (2015) start from the work Nishiyama et al. (2011), where the process governing $\{y_t\}_{t \in T}$ takes the following form

$$
y_t = \gamma(Y_{t-1}) + \varrho(X_{t-1}) + \zeta_t,
$$

with $\zeta_t$ being a white noise process, whereas $\gamma(\cdot)$ and $\varrho(\cdot)$ are unknown functions satisfying condi-
tions ensuring stationarity of $y_t$.

Balcilar et al. (2015) noticed that the specification in (7) 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 $\varphi(\cdot)$ is a general nonlinear function. Therefore, the model allows deriving a general test for quantile causality in variance, where the impact does not need to come from squared values of $x_{t-1}$. Notably, the test could detect an impact from levels or non-linear transformations of $x_{t-1}$ to the squared values of $y_t$, where the squared of $y_t$ is just a proxy of the conditional variance of $y_t$. Clearly, this corresponds to an implicit assumption of heteroskedasticity with the variance driven by (transformed) lagged values of $x_t$ (and $y_t$).

To introduce a test for quantile causality in variance, Balcilar et al. (2015) reformulate Equation (7) as follows:

$$
\begin{align*}
H_0 : P[F_{y_t^2|Z_{t-1}}(Q_{\tau}(Y_{t-1})|Z_{t-1}) = \tau] = 1 \\
H_1 : P[F_{y_t^2|Z_{t-1}}(Q_{\tau}(Y_{t-1})|Z_{t-1}) = \tau] < 1.
\end{align*}
$$

(8)

In order to test the hypotheses in (8), Balcilar et al. (2015) proposed using the test statistic $\hat{J}_T$, replacing $y_t$ by $y_t^2$, and preserving the same asymptotic distribution.

In our empirical analyses, for both quantile causality in mean and variance, we computed $\hat{J}_T$ making use of the Gaussian kernel for both $K(\cdot)$ and $C(\cdot)$, with bandwidths obtained through the least squares cross-validation method (Jeong et al., 2012; Balcilar et al., 2015).

### 2.2 Quantiles and density forecasting

While Section 2.1 focuses on the presence of quantile causality, the present Section introduces methods aiming at the forecasting implications of EMU and EPU. As mentioned in the introduction, one of the research questions points at the evaluation of the potentially different impact of the two uncertainty indexes. A forecasting exercise allows for a direct comparison of the two, which can be introduced jointly in a model, allowing for testing on their statistical impact as well as for their forecasting impact. We thus introduce in the forecasting exercise both EMU and EPU, whose joint impact could provide important implications given that they quantify two different sources of uncertainty. In particular, we aim to forecast both the conditional quantiles and distributions of $y_t$ and $y_t^2$ taking into account the information associated with the two indexes.

The first step consists in estimating the conditional quantiles and, for this purpose, we make use of the quantile regression approach introduced by Koenker and Bassett (1978). We first remind that $1_{\{\cdot\}}$ is the indicator function taking value 1 if the condition in $\{\cdot\}$ is true, 0 otherwise. The approach
introduced by Koenker and Bassett (1978) makes use of the asymmetric loss function
\[ \rho_\tau(\varepsilon) = \varepsilon [\tau - 1_{\{\varepsilon < 0\}}]. \]  

Starting from the case of \( y_t \) and given \( W_{t-1} = (y_{t-1}, ..., y_{t-p}, x_{1,t-1}, ..., x_{1,t-q}, x_{2,t-1}, ..., x_{2,t-r}) \), for \((p,q,r) > 1\), Koenker and Bassett (1978) showed that the minimizer, \( Q_\tau(y_t|W_{t-1}) \), of the expected loss
\[ E[\rho_\tau(y_t - Q_\tau(y_t|W_{t-1})] \]  
satisfies \( F_Y(Q_\tau(y_t|W_{t-1})) - \tau = 0 \), where \( Q_\tau(y_t|W_{t-1}) \), the conditional \( \tau \)-th quantile of \( y_t \), is equal to
\[ Q_\tau(y_t|W_{t-1}) = \alpha_0(\tau) + \beta_1(\tau)y_{t-1} + ... + \beta_p(\tau)y_{t-p} + \delta_1(\tau)x_{1,t-1} + ... + \delta_q(\tau)x_{1,t-q} + \lambda_1(\tau)x_{2,t-1} + ... + \lambda_r(\tau)x_{2,t-r}. \]  

The unknown parameters in Equation (11) are estimated by minimizing equation (10). Then, with a horizon of one period ahead, the forecast of the \( \tau \)-th conditional quantile of \( y_t \) is computed as
\[ \hat{Q}_\tau(y_{t+1}|W_t) = \hat{\alpha}_0(\tau) + \hat{\beta}_1(\tau)y_t + ... + \hat{\beta}_p(\tau)y_{t-p+1} + \hat{\delta}_1(\tau)x_{1,t} + ... + \hat{\delta}_q(\tau)x_{1,t-q+1} + \hat{\lambda}_1(\tau)x_{2,t} + ... + \hat{\lambda}_r(\tau)x_{2,t-r+1}. \]  

After obtaining a grid of forecasted quantiles, computed at different \( \tau \) values, the second step consists in forecasting the conditional distribution of the oil returns. The standard quantile regression approach allows estimating individual quantiles, but it does not guarantee their coherence, i.e. their increasing monotonicity in \( \tau \in (0, 1) \). For instance, it might occur that the predicted 95-th percentile of the response variable is lower than the 90-th percentile. If quantiles cross, corrections must be applied in order to obtain a valid conditional distribution of volatility. For instance, in order to cope with the crossing problem, Koenker (1984) applied parallel quantile planes, whereas Bondell et al. (2010) estimated the quantile regression coefficients with a constrained optimization method.

Here we follow a different approach, proposed by Zhao (2011). Given a collection of \( \vartheta \) predicted conditional quantiles \( \hat{Q}_{\tau_1}(y_{t+1}|W_t), ..., \hat{Q}_{\tau_{\vartheta}}(y_{t+1}|W_t) \), for \( 0 < \tau_j < \tau_{j+1} < 1, \ j = 1, ..., \vartheta - 1 \), we first rearrange them into ascending order, by making use of the quantile bootstrap method proposed by Chernozhukov et al. (2010). Then, starting from the rearranged quantiles, denoted by \( \hat{Q}_{\tau_1}^*(y_{t+1}|W_t), ..., \hat{Q}_{\tau_{\vartheta}}^*(y_{t+1}|W_t) \), we estimate the entire conditional distribution with a nonparametric kernel method. The predicted density equals
\[ \hat{f}_{y_{t+1}|W_t}(y^*|W_t) = \frac{1}{\vartheta h_{\vartheta}} \sum_{i=1}^{\vartheta} K_h \left( \frac{y^* - \hat{Q}_{\tau_i}^*(y_{t+1}|W_t)}{h_{\vartheta}} \right), \]  

(13)
where \( y^* \) are evenly interpolated points that generates the support of the estimated distribution, \( h_\theta \) is the bandwidth, \( K_e(\cdot) \) is the kernel function, and \( \hat{f}(\cdot|\mathcal{W}_t) \equiv \hat{f}_{y_{t+1}|\mathcal{W}_t}(\cdot|\mathcal{W}_{t-1}) \) is the one-period ahead forecasted density, given the information set available in \( t \). Following Gaglianone and Lima (2012), we use as \( K_e(\cdot) \) the Epanechnikov kernel.

Given \( \mathcal{W}_{2,t-1} \equiv (y_{t-1}^2, ..., y_{t-p}^2, x_{1,t-1}, ..., x_{1,t-q}, x_{2,t-1}, ..., x_{2,t-r}) \), for \((p, q, r) > 1\), the conditional distribution of \( y_t^2 \), denoted as \( \hat{f}_{y_{t+1}|\mathcal{W}_{2,t}}(\cdot|\mathcal{W}_{2,t}) \), is obtained by applying the same methodology described above, by replacing \( Q_{\tau}^*(y_{t+1}|\mathcal{W}_t) \) by \( Q_{\tau}^*(y_{t+1}^2|\mathcal{W}_{2,t}) \). Specifically, \( Q_{\tau}^*(y_{t+1}^2|\mathcal{W}_{2,t}) \) is the conditional \( \tau \)-th quantile of \( y_t^2 \), adjusted for the crossing quantiles issue, arising from the original one \( Q_{\tau}(y_{t+1}^2|\mathcal{W}_{2,t}) \). We estimated the latter as

\[
\hat{Q}_{\tau}(y_{t+1}^2|\mathcal{W}_{2,t}) = \hat{\alpha}_0(\tau) + \hat{\beta}_1(\tau)y_t^2 + ... + \hat{\beta}_p(\tau)y_{t-p+1}^2 + \hat{\delta}_1(\tau)x_{1,t} + ... + \hat{\delta}_q(\tau)x_{1,t-q+1} + \hat{\lambda}_1(\tau)x_{2,t} + ... + \hat{\lambda}_r(\tau)x_{2,t-r+1}.
\]

(14)

We compute the coefficients standard errors through the bootstrap method (Efron, 1979), whose advantages are well-known: it assumes no particular distribution of the errors, it is not based on asymptotic model properties and it is available regardless of the statistic of interest’s complexity. Among all the available bootstrapping methods, we make use of the \( xy \)-pair method (Kocherginsky, 2003), whose advantages for quantile regression problems are highlighted in Davino et al. (2014).

### 2.3 Evaluation of the predictive accuracy

We evaluate the predictive accuracy of the method described in Section 2.2 by using five testing approaches, introduced, respectively, by Berkowitz (2001), Diebold and Mariano (2002), Amisano and Giacomini (2007), Diks et al. (2011), Gneiting and Ranjan (2011). In the following, we give the main details about the five tests below, focusing on the conditional quantiles and distribution of \( y_t \). The same methodology applies to \( y_t^2 \) with \( Q_{\tau}^*(y_t^2|\mathcal{W}_{2,t}) \) replacing \( Q_{\tau}(y_t^2|\mathcal{W}_t) \). We also remind that the forecast evaluation takes as input a collection of one-step-ahead forecasts.

As regards the Berkowitz (2001) test, we first compute the variable

\[
\psi_{t+1} = \int_{-\infty}^{y_{t+1}} \hat{f}(u|\mathcal{W}_t)du = \hat{F}(y_{t+1}|\mathcal{W}_t),
\]

(15)

where \( \hat{F}(y_{t+1}|\mathcal{W}_t) \) is the distribution function corresponding to the density \( \hat{f}(y_{t+1}|\mathcal{W}_t) \); \( \psi_{t+1} \) is computed sequentially \( M \) times, where \( M < T \) is the number of periods included in the interval spanning the forecasting evaluation.

Rosenblatt (1952) showed that, if the model is correctly specified, \( \psi_{t+1} \) is i.i.d. and uniformly distributed on \((0, 1)\); that result holds regardless of the \( y_t \) distribution, even if \( \hat{F}(\cdot|\mathcal{W}_t) \) changes over
time. Berkowitz (2001) observed that, if $\psi_{t+1} \sim U(0, 1)$, then

$$z_{t+1} = \Phi^{-1}(\psi_{t+1}) \sim N(0, 1), \quad (16)$$

where $\Phi^{-1}(\cdot)$ denotes the inverse of the standard normal distribution function.

Given that, under correct model specification, $z_{t+1}$ should be independent and identically distributed as standard normal, an alternative hypothesis is that the mean and the variance differ from 0 and 1, respectively, with a first-order autoregressive structure. In particular, Berkowitz (2001) considered the model

$$z_{t+1} - \mu_b = \rho_b (z_t - \mu_b) + e_{t+1} \quad (17)$$

to test the null hypothesis $H_0 : \mu_b = 0, \rho_b = 0, \operatorname{var}(e_{t+1}) = \sigma_b^2 = 1$, which corresponds to the appropriate specification of the density forecasting model. The test built on (17) is based on the likelihood-ratio statistic

$$LR_b = -2 \left[ L_b(0, 1, 0; z_{t+1}) - L_b(\hat{\mu}_b, \hat{\sigma}_b, \hat{\rho}_b; z_{t+1}) \right], \quad (18)$$

where $L_b(\hat{\mu}_b, \hat{\sigma}_b, \hat{\rho}_b; z_{t+1})$ is the likelihood function associated with Equation (17) and computed from the maximum-likelihood estimates of the unknown parameters $\mu_b, \sigma_b$ and $\rho_b$. Under the null hypothesis $H_0$, the test statistic is distributed as $\chi^2(3)$.

First of all, we use the Berkowitz test for an absolute assessment of the density forecasts recovered from (13). Then, we also implement the test on a restricted model, i.e. the one which uses in Equation (11) just $y_{t-1} \equiv (y_{t-1}, ..., y_{t-p})$ as predictors; we denote by $\hat{f}(\cdot|Y_t)$ the density we forecast on the basis of the restricted model. Hence, we can assess the joint contribution of EPU and EMU in predicting the distribution of the oil returns by comparing the $LR_b$ values arising from the unrestricted and the restricted models. We also evaluate the contribution of each uncertainty index separately, by adding to the restricted model just the lagged values of $x_{1,t}$ when we focus on EPU, or the lagged values of $x_{2,t}$, when we consider EMU.

The approach proposed by Berkowitz (2001) evaluates the goodness of a specific sequence of density forecasts, relative to the unknown data-generating process. However, given a certain model, the Berkowitz test has power only for misspecifications of the first two moments, but in practice, that model could be misspecified at higher-order moments. In that case, a valid solution consists in comparing density forecasts, i.e. performing a relative comparison given a specific measure of accuracy. Hence, in addition to the approach proposed by Berkowitz (2001), we also consider the tests introduced by Diebold and Mariano (2002), Amisano and Giacomini (2007), Diks et al. (2011), Gneiting and Ranjan (2011).

We implement the test developed by Diebold and Mariano (2002) on the basis of the losses gener-
ated by the unrestricted and the restricted models, denoted by $L_{\tau,t+1}(y_{t+1}|W_t)$ and $L_{\tau,t+1}(y_{t+1}|Y_t)$, respectively. Among the various loss functions adopted in the literature, following Giglio et al. (2012), we make use of those defined as follows:

$$L_{\tau,t+1}(y_{t+1}|W_t) = \left( \tau - 1 \{y_{t+1} - Q_{\tau}^*(y_{t+1}|W_t) < 0 \} \right) [y_{t+1} - Q_{\tau}^*(y_{t+1}|W_t)],$$

(19)

$$L_{\tau,t+1}(y_{t+1}|Y_t) = \left( \tau - 1 \{y_{t+1} - Q_{\tau}^*(y_{t+1}|Y_t) < 0 \} \right) [y_{t+1} - Q_{\tau}^*(y_{t+1}|Y_t)].$$

(20)

Given the loss differential,

$$d_{DM,\tau,t+1} = L_{\tau,t+1}(y_{t+1}|W_t) - L_{\tau,t+1}(y_{t+1}|Y_t),$$

(21)

which we evaluate for all the periods included in $[t+1, t+M]$, we compute its average value, denoted by $d_{DM,\tau}$.

We are interested in testing the null hypothesis $H_0 : E[d_{DM,\tau}] = 0$ against the alternative $H_1 : E[d_{DM,\tau}] \neq 0$. For that purpose, Diebold and Mariano (2002) proposed the test statistic:

$$DM_\tau = \frac{d_{DM,\tau}}{\hat{\sigma}_{DM}/\sqrt{M}},$$

(22)

where $\hat{\sigma}^2_{DM}$ is a consistent estimate of $\sigma^2_{DM} = Var \left( \sqrt{M} d_{DM,\tau} \right)$, the asymptotic (long-run) variance. Diebold and Mariano (2002) showed that, under the null hypothesis of equal predictive accuracy, $DM_\tau \xrightarrow{d} N(0,1)$; in case the null hypothesis is rejected and the (22) takes negative values, we have evidence for the unrestricted model having better performance. To evaluate the $DM_\tau$ test statistic, we focus on selected quantiles, setting $\tau = \{0.05, 0.5, 0.95\}$. In this way, we do not evaluate the entire density forecast, but only the impact of uncertainty indexes on selected quantiles.

The next tests focus on the entire density forecast, thus allowing for a much broader evaluation of the uncertainty indexes relevance. The approach introduced by Amisano and Giacomini (2007) compares two different competing models on the basis of their log-scores. In particular, the log-scores of the unrestricted and the restricted models are denoted by $\log \left( \hat{f}(y_{t+1}|W_t) \right)$ and $\log \left( \hat{f}(y_{t+1}|Y_t) \right)$, respectively. Given a sequences of density forecasts, it is possible to compute the quantity defined as

$$WLR_{t+1} = w(y_{t+1}^{qt}) \left[ \log \left( \hat{f}(y_{t+1}|W_t) \right) - \log \left( \hat{f}(y_{t+1}|Y_t) \right) \right],$$

(23)

where $y_{t+1}^{qt}$ is the standardized oil return in $t+1$, whereas $w(y_{t+1}^{qt})$ is the weight the forecaster arbitrarily chooses to emphasize particular regions of the distribution’s support.

After computing $WLR_{t+1}$ for the $M$ periods included in the interval spanning the forecast eval-
uation, we evaluate its mean, which we denoted by $\overline{\text{WLR}}$. In order to test the null hypothesis of equal performance, that is, $H_0 : E[\text{WLR}] = 0$, against the alternative of a different predictive ability $H_1 : E[\text{WLR}] \neq 0$, Amisano and Giacomini (2007) suggest the use of a weighted likelihood ratio test:

$$ AG = \frac{\text{WLR}}{\hat{\sigma}_{AG}/\sqrt{M}}, $$

(24)

where $\hat{\sigma}_{AG}^2$ is a heteroskedasticity- and autocorrelation-consistent (HAC) Newey and West (1987) estimator of $\sigma_{AG}^2 = \text{Var}\left(\sqrt{M} \text{WLR}\right)$. $AG$ is positive in case of a better performance of the unrestricted model, otherwise it takes negative values. Amisano and Giacomini (2007) showed that, under the null hypothesis, $AG \xrightarrow{d} \mathcal{N}(0, 1)$.

We applied the Amisano and Giacomini (2007) test by using four designs for the weights entering Equation (23), in order to verify how the results change according to the particular regions of the distribution’s support on which we are focusing. We set $w_{CE}(y_{t+1}) = \phi(y_{t+1})$ to give an higher weight to the center of the distribution, $w_{RT}(y_{t+1}) = \Phi(y_{t+1})$ when we focus more on the right tail, $w_{LT}(y_{t+1}) = 1 - \Phi(y_{t+1})$ for the left tail, and $w_{NW}(y_{t+1}) = 1$ when giving equal importance to the entire support.

As noticed by Diks et al. (2011) and Gneiting and Ranjan (2011), the weighted logarithmic scoring rule as in Amisano and Giacomini (2007) favors density forecasts with more probability mass in the region of interest and, as a result, the resulting test of equal predictive ability is biased toward such density forecasts; hence, they proposed different scores to solve this shortcoming of the Amisano and Giacomini (2007) test.

In the case of the unrestricted model, the score introduced by Diks et al. (2011) is defined as

$$ S_{csl}^\text{est}(y_{t+1}|\text{W}_t) = w_{csl,t}(y_{t+1}) \log \hat{f}(y_{t+1}|\text{W}_t) + (1 - w_{csl,t}(y_{t+1})) \log \left[1 - \int w_{csl,t}(s) \hat{f}(s|\text{W}_t) ds\right], $$

(25)

where $w_{csl,t}(\cdot)$ is the weighting function, by which we focus on the density’s region of interest, whereas the second addend in (25) avoids the mistake of attaching comparable scores to density forecasts that have similar tail shapes but may have completely different tail probabilities (Diks et al., 2011).

Let $\bar{y}_1$ and $\bar{y}_3$ be the in-sample first and third quartile of $y_t$, respectively, we set $w_{csl,t}(y_{t+1}) = 1_{\{y_{t+1} \leq \bar{y}_1\}}$ when we focus on the left tail, $w_{csl,t}(y_{t+1}) = 1_{\{\bar{y}_1 \leq y_{t+1} \leq \bar{y}_3\}}$ when we place the attention on the center of the distribution, $w_{csl,t}(y_{t+1}) = 1_{\{y_{t+1} \geq \bar{y}_3\}}$ when we consider the right tail.

Similarly, we denote the score function of the restricted model as $S_{csl}^\text{est}(y_{t+1}|\text{Y}_t)$, obtained by replacing $\text{W}_t$ by $\text{Y}_t$ in (25).

Let $S_{csl}$ be the mean of the differences $S_{csl}^\text{est}(y_{t+1}|\text{W}_t) - S_{csl}^\text{est}(y_{t+1}|\text{Y}_t)$, computed for all the periods.

\footnote{Note that $\phi(.)$ and $\Phi(.)$ denote the standard normal density function and the standard normal distribution function, respectively.}
including the interval $[t+1, t+M]$. The test statistic proposed by Diks et al. (2011) equals:

$$DPD = \frac{\tilde{S}^{cal}}{\tilde{\sigma}_{DPD}/\sqrt{M}}, \quad (26)$$

where $\hat{\sigma}_{DPD}^2$ is a heteroskedasticity- and autocorrelation-consistent (HAC) Newey and West (1987) estimator of $\sigma^2_{DPD} = \text{Var} \left( \sqrt{M} \tilde{S}^{cal} \right)$. We have evidence of a better/worse performance of the unrestricted model when $DPD$ takes positive/negative values. Diks et al. (2011) showed that, under the null hypothesis of equal performance, $DPD \xrightarrow{d} \mathcal{N}(0, 1)$.

Finally, again focusing on the unrestricted model, the score proposed by Gneiting and Ranjan (2011) is defined as follows:

$$S^{gr}(y_{t+1}|W_t) = \frac{1}{I-1} \sum_{i=1}^{I} w(\tau_i) QS_{\tau_i} \left[ \hat{F}^{-1}(\tau_i|W_t), y_{t+1} \right], \quad (27)$$

where $\tau_i = i/I$ and

$$QS_{\tau_i} \left[ \hat{F}^{-1}(\tau_i|W_t), y_{t+1} \right] = 2 \left[ 1_{\{y_{t+1} < \hat{F}^{-1}(\tau_i|W_t)\}} - \tau_i \right] (\hat{F}^{-1}(\tau_i|W_t) - y_{t+1}). \quad (28)$$

It is interesting to observe that the quantity defined in (28) is similar to the one in (19); nevertheless, the loss given in (27) is more informative than $L_{\tau,t+1}(y_{t+1}|W_t)$, since it is equal to the weighted average of several $QS_{\tau_i} \left[ \hat{F}^{-1}(\tau_i|W_t), y_{t+1} \right]$ values computed for a sufficiently large grid of probabilities levels.

As for the weight function, as suggested by Gneiting and Ranjan (2011), we set $w(\tau_i) = \tau_i (1 - \tau_i)$, $w(\tau_i) = \tau_i^2$, $w(\tau_i) = (1 - \tau_i)^2$ to assign greater importance to the center, the right tail and the left tail of the distribution, respectively. Similarly, we denote the score arising from the restricted model as $S^{gr}(y_{t+1}|Y_t)$; we stress we obtain the score by replacing $W_t$ by $Y_t$ in (28). Let $\bar{S}^{gr}$ be the average value of the differences $S^{gr}(y_{t+1}|W_t) - S^{gr}(y_{t+1}|Y_t)$ computed for all the periods included in $[t+1, t+M]$, the null hypothesis of equal performance is tested through the statistic

$$GR = \frac{\bar{S}^{gr}}{\hat{\sigma}_{GR}/\sqrt{M}}, \quad (29)$$

where $\hat{\sigma}_{GR}^2$ is a heteroskedasticity- and autocorrelation-consistent (HAC) Newey and West (1987) estimator of $\sigma^2_{GR} = \text{Var} \left( \sqrt{M} \bar{S}^{gr} \right)$. We have evidence of a better/worse performance of the unrestricted model when $GR$ takes negative/positive values. Gneiting and Ranjan (2011) showed that, under the null hypothesis, $GR \xrightarrow{d} \mathcal{N}(0, 1)$. 

12
3 Empirical set-up

3.1 Data description

In our analyses we make use of three series: the oil prices and two uncertainty indexes, EPU and EMU. The series are sampled at daily frequency and cover the period between January 2, 1986 and April 23, 2015, for a total of 7646 days.

We denote by \( \{y_t\}_{t \in T} \) the series of the oil returns, that is \( y_t = \log(oil_t) - \log(oil_{t-1}) \), where \( oil_t \) is the spot price of the West Texas Intermediate (WTI) crude oil at day \( t \).\(^5\) \( \{oil_t\}_{t \in T} \) is not stationary: both the augmented Dickey and Fuller (1981) and the Phillips and Perron (1988) tests don’t reject the null hypothesis of unit root with p-values of 0.2623 and 0.2112, respectively; differently, the p-values of the two tests are less than 0.01 for both \( \{y_t\}_{t \in T} \) and \( \{y^2_t\}_{t \in T} \).

\( EPU \) and \( EMU \) are two indices measuring the US economic policy and equity market uncertainty.\(^6\) \( EPU \) is built from newspaper archives of the Access World News’s NewsBank service, by restricting the attention on United States and taking into account the number of articles containing at least one of the terms belonging to 3 sets. The first set is “economic/economy”, the second is “uncertain/uncertainty” and the third set is “legislation/deficit/regulation/congress/federal reserve/white house”. Using the same news source, \( EMU \) is built from articles containing the terms previously mentioned and one or more of the following: “equity market/equity price/stock market”. From \( EPU \) and \( EMU \) we compute \( \{x_{1,t}\}_{t \in T} \) and \( \{x_{2,t}\}_{t \in T} \), which are not affected by unit root: in both the cases the p-values of the augmented Dickey and Fuller (1981) and of the Phillips and Perron (1988) tests are less than 0.01.

Table 1: Descriptive statistics.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>St. Deviation</th>
<th>Min</th>
<th>Max</th>
<th>skewness</th>
<th>kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>( y_t )</td>
<td>0.0001</td>
<td>0.0248</td>
<td>-0.4069</td>
<td>0.1924</td>
<td>-0.7639</td>
<td>18.3329</td>
</tr>
<tr>
<td>( y^2_t )</td>
<td>0.0006</td>
<td>0.0026</td>
<td>0.0000</td>
<td>0.1655</td>
<td>37.9947</td>
<td>2294.8710</td>
</tr>
<tr>
<td>( x_{1,t} )</td>
<td>4.3665</td>
<td>0.6776</td>
<td>1.2185</td>
<td>6.5780</td>
<td>-0.2679</td>
<td>3.2726</td>
</tr>
<tr>
<td>( x_{2,t} )</td>
<td>3.8459</td>
<td>1.0575</td>
<td>1.5688</td>
<td>7.8655</td>
<td>0.2718</td>
<td>2.7157</td>
</tr>
</tbody>
</table>

The table reports some descriptive statistics computed for \( y_t \), \( y^2_t \), \( x_{1,t} \) and \( x_{2,t} \). From left to right we report the mean, the standard deviation, the minimum and maximum values, the skewness and the kurtosis indices.

We report in Table 1 some descriptive statistics computed for the variables above described. \( y_t \) and \( y^2_t \) have average values close to zero, with standard deviations equal to 0.0248 and 0.0026, respectively; \( y_t \) ranges from -0.4069 to 0.1924 and its distribution is affected by negative skewness and leptokurtosis. \( y^2_t \) has strong positive skewness and leptokurtosis, due to the presence of relevant extreme values in its right tail. The uncertainty indexes, \( x_{1,t} \) and \( x_{2,t} \), are centered around 4.366 and 3.8459, with standard deviations equal to 0.6776 and 1.0575, respectively. Their distributions are slightly skewed, quite mesocurtic and affected by the presence of a few extreme values in the tails. The explorative analysis

\(^5\) The series of the oil prices is recovered from Thomson Reuters Datastream.
\(^6\) The data and the details about EPU and EMU are available on http://www.policyuncertainty.com/.
highlights the presence of extreme values for the variables of interest, mainly $y_t$ and $y_{t}^{2}$, suggesting the wisdom of using the quantile regression (Koenker and Bassett, 1978) in the forecasting exercise, rather than the ordinary least squares approach, because the latter does not guarantee robust results in the presence of outliers.

We now move to the most relevant empirical analyses, and investigate the ability of $x_{1,t} = \log(EPU_t)$ and $x_{2,t} = \log(EMU_t)$ in causing and predicting both the oil returns ($y_t$) and the squared oil returns ($y_{t}^{2}$) series, the latter being a measure of risk or dispersion of oil returns. We first describe the approach we follow in deriving the out-of-sample density forecasts.

### 3.2 Dynamic analysis and rolling window procedure

As noticed by Balcilar et al. (2015), the relationships among $y_t$ or $y_{t}^{2}$ and the uncertainty indices are not stable over time. They applied the Bai and Perron (2003) test, detecting the presence of multiple structural breaks in the oil returns series for the EPU- and EMU-based VARs. Here, we follow a different approach, by implementing the $DQ$ and the $SQ$ tests introduced by Qu (2008), which reveal structural changes with unknown timing in regression quantiles. Following Tillmann and Wolters (2015), whose study focuses on the US inflation persistence, we proceed in two stages.

First, we use the $DQ$ test in order to capture possible changes in the entire conditional distribution of the response variable. Given that we do not have any prior information as to which part of the conditional distribution is affected by breaks, we take into account a large range of quantiles levels, namely $\tau = \{0.05, 0.1, 0.15, \ldots, 0.95\}$.

By setting $p = q = r = 2$ in Models (12)-(14), given $\xi(\tau) = [\alpha_0(\tau), \beta_1(\tau), \beta_2(\tau), \delta_1(\tau), \delta_2(\tau), \lambda_1(\tau), \lambda_2(\tau)]$ and $1 \leq T_1 < T^* < T_2 \leq T$, the hypotheses of the $DQ$ tests are defined as follows:

$$
H_0 : \xi_t(\tau) = \xi(\tau), \text{ for all } t \text{ and for all } \tau \in \{0.05, 0.1, 0.15, \ldots, 0.95\}
$$

$$
H_1 : \xi_t(\tau) = \begin{cases} 
\xi_1(\tau), & \text{for } t = T_1, \ldots, T^* \\
\xi_2(\tau), & \text{for } t = T^*, \ldots, T_2 
\end{cases}, \text{ for some } \tau \in \{0.05, 0.1, 0.15, \ldots, 0.95\}.
$$

In a second step, we implement, at the dates where the null hypothesis of the $DQ$ test is rejected at the level of 0.01, the $SQ$ test; in this way we detect structural changes in prespecified quantiles, in order to identify the specific regions of the distribution affected by breaks; for simplicity, we implemented the $SQ$ test at three quantiles levels, i.e. $\tau = \{0.1, 0.5, 0.9\}$. In the present work, the hypotheses of

---

7We were also able to detect four (18/01/1991, 26/03/2003, 02/12/2008, and 05/11/2011) and five (18/02/1999, 24/03/2003, 31/05/2007, 11/12/2008 and 05/11/2011) breaks with EPU and EMU being the independent variables respectively, in relation to oil returns.
the SQ tests are defined as follows:

\[
\begin{align*}
H_0 : \xi_t(\tau) &= \xi(\tau), \text{ for all } t \text{ and for a given } \tau \in \{0.1, 0.5, 0.9\} \\
H_1 : \xi_t(\tau) &= \begin{cases} 
\xi_1(\tau), & \text{for } t = T_1, \ldots, T^* \\
\xi_2(\tau), & \text{for } t = T^*, \ldots, T_2 \end{cases}, \text{ for a given } \tau \in \{0.1, 0.5, 0.9\}.
\end{align*}
\]

(31)

The tests proposed by Qu (2008) are subgradient and have good properties also in small samples. The tables containing the critical values of the DQ and the SQ tests are available in Qu (2008).

The output of the DQ test, applied to the conditional quantiles and distribution of \(y_t\) is given in the left panel of Table 2. The number of breaks, detected at the level of 0.01, is equal to 7. The results of the SQ test are given in the right panel of Table 2; here, we can see that the breaks mainly affect the extreme conditional quantiles of \(y_t\), rather than the central ones.

Table 2: Structural breaks in the conditional distribution and quantiles of \(y_t\).

<table>
<thead>
<tr>
<th>Structural breaks in the conditional distribution</th>
<th>Structural breaks at specific quantiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dates of breaks</td>
<td>(DQ) \quad \text{(SQ(\tau=0.1))} \quad \text{(SQ(\tau=0.5))} \quad \text{(SQ(\tau=0.9))}</td>
</tr>
<tr>
<td>20/03/1987</td>
<td>1.0733</td>
</tr>
<tr>
<td>11/05/1989</td>
<td>1.0852</td>
</tr>
<tr>
<td>21/09/1990</td>
<td>1.0644</td>
</tr>
<tr>
<td>05/11/1991</td>
<td>1.0826</td>
</tr>
<tr>
<td>04/09/2000</td>
<td>1.0522</td>
</tr>
<tr>
<td>16/08/2013</td>
<td>1.0598</td>
</tr>
<tr>
<td>27/01/2015</td>
<td>1.1041</td>
</tr>
</tbody>
</table>

The table reports the output of the DQ and the SQ tests, introduced by Qu (2008). The former detects the presence of structural breaks in the conditional distribution of \(y_t\) at the level of 0.01, whereas the latter detects the presence of structural breaks at specific quantiles, namely at \(\tau = \{0.1, 0.5, 0.9\}\); *, ** and *** refer, respectively, to the 10%, 5% and 1% significance level.

Similarly, we show the results of the two tests, arising from the estimation of the \(y_t^2\) conditional quantiles and distribution, in Table 3. The number of breaks is equal to 13 and the structural changes mainly occur at medium-high levels of \(\tau\).

The results discussed above highlight the presence of structural breaks over time and, as a result, the conclusions drawn from the full sample analysis might not be consistent. In order to capture the dynamics in the relations among the variables of interest, we differ from Balcilar et al. (2015) by implementing a rolling window procedure for causality testing, model estimation and forecast computation. The window used for the estimation of the model has a width of 500 observations. Moreover, to make a balance between flexibility, efficiency, and computational burden, we re-estimated the model with step of 5 days. In details, and focusing on causality testing at quantiles, the first window we consider includes the observations recorded between the first and the 500-th day of the sample. At time \(t = 500\), we compute, for the first time, \(\hat{J}_T\) at different quantiles levels, with \(\tau\) ranging from 0.05 to 0.95 and step of 0.05, for a total of 19 \(\hat{J}_T\) values.

At \(t = 500\), we also estimate, for the first time, the parameters of the models defined, respectively,
Table 3: Structural breaks in the conditional distribution and quantiles of $y_{2t}$.

<table>
<thead>
<tr>
<th>Dates of breaks</th>
<th>DQ</th>
<th>SQ ($\tau = 0.1$)</th>
<th>SQ ($\tau = 0.5$)</th>
<th>SQ ($\tau = 0.9$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>08/09/1986</td>
<td>1.0669</td>
<td>1.4449</td>
<td>1.7797 **</td>
<td>1.4096</td>
</tr>
<tr>
<td>08/12/1988</td>
<td>1.0630</td>
<td>1.3808</td>
<td>2.0024 ***</td>
<td>1.5430 **</td>
</tr>
<tr>
<td>09/10/1989</td>
<td>1.0599</td>
<td>1.2835</td>
<td>1.7168 **</td>
<td>1.5511 *</td>
</tr>
<tr>
<td>26/09/1990</td>
<td>1.1024</td>
<td>0.9704</td>
<td>1.3059</td>
<td>2.1648 ***</td>
</tr>
<tr>
<td>03/07/1991</td>
<td>1.0663</td>
<td>0.9263</td>
<td>2.0506 ***</td>
<td>1.9799 ***</td>
</tr>
<tr>
<td>27/01/1993</td>
<td>1.1142</td>
<td>0.9083</td>
<td>2.0441 ***</td>
<td>1.4157</td>
</tr>
<tr>
<td>16/03/1994</td>
<td>1.0746</td>
<td>1.4506</td>
<td>1.3358</td>
<td>1.7598 **</td>
</tr>
<tr>
<td>13/12/1994</td>
<td>1.0769</td>
<td>0.8698</td>
<td>1.8874 ***</td>
<td>0.8569</td>
</tr>
<tr>
<td>24/06/1996</td>
<td>1.0593</td>
<td>1.5555 *</td>
<td>1.7563 **</td>
<td>1.8971 ***</td>
</tr>
<tr>
<td>26/03/1999</td>
<td>1.0827</td>
<td>1.2611</td>
<td>1.0853</td>
<td>1.6581 *</td>
</tr>
<tr>
<td>18/05/2009</td>
<td>1.0926</td>
<td>1.3832</td>
<td>1.8667 ***</td>
<td>2.6657</td>
</tr>
<tr>
<td>04/12/2012</td>
<td>1.1001</td>
<td>0.9915</td>
<td>1.4124</td>
<td>1.4504</td>
</tr>
<tr>
<td>24/02/2015</td>
<td>1.0721</td>
<td>1.6582 **</td>
<td>1.4334</td>
<td>1.5175 *</td>
</tr>
</tbody>
</table>

The table reports the output of the DQ and the SQ tests, introduced by Qu (2008). The former detects the presence of structural breaks in the conditional distribution of $y_{2t}$ at the level of 0.01, whereas the latter detects the presence of structural breaks at specific quantiles, namely at $\tau = \{0.1, 0.5, 0.9\}$; *, ** and *** refer, respectively, to the 10%, 5% and 1% significance level.

in (12) and (14), by setting $\tau$ from 0.01 to 0.99, with step of 0.01, to obtain quantiles vectors of length 99. The finer grid of quantiles used in the forecasting exercise, with respect to the causality analysis, is due to the need of estimating with adequate precision the conditional distributions of $y_t$ and $y_{2t}^2$. Given the parameter estimates obtained at time $t = 500$, we compute the forecasts of the conditional quantiles and distributions of $y_t$ and $y_{2t}^2$ for $t = 501, ..., 505$. Note we are not making a 5-step-ahead forecast, but simply fix the model parameters for 5 days, and compute five one-step-ahead forecasts. For instance, to recover the quantile forecasts we multiply the values of the predictors observed in $t = 500, ..., 504$ by the coefficients estimated in $t = 500$.

The second window includes the observations between the 6-th and the 505-th day. Hence, at $t = 505$, we compute for the second time, updating the previous output obtained in $t = 500$, both $\hat{J}_T$ and the estimated parameters by which we forecast, for $t = 506, ..., 510$, the conditional quantiles and distributions of $y_t$ and $y_{2t}^2$. The procedure goes on until the entire dataset is completely exploited.

As for the implementation of the tests proposed by Berkowitz (2001), Diebold and Mariano (2002), Amisano and Giacomini (2007), Diks et al. (2011) and Gneiting and Ranjan (2011), described in Section 2.3, starting from $t + 1 = 501$, we compare the forecasts formulated in $t$ with the out-of-sample observations $y_{t+1}$ and $y_{2t+1}^2$. Therefore, on the basis of those comparisons, we compute the quantity in (16) along with the scores characterizing each of the tests defined, respectively, in (22), (24), (26), (29). In our analysis, the forecasting evaluation is carried out on rolling intervals consisting of $M = 500$ periods. Therefore, in $t = 1000$, we compute for the first time the five test statistics mentioned above. By updating the (16) and the scores by one period ahead, we compute the statistics for the second time in $t = 1001$, and the procedure goes on until the entire dataset is completely exploited. Note that we make use of two windows: the first refers to the model estimation, while the latter defines the range.
over which we evaluate the density forecast performances of the restricted and unrestricted models.

In applying the test introduced by Jeong et al. (2012), in Balcilar et al. (2015) the lag order \( q \) is determined on the basis of the Schwarz Information Criterion computed on the VAR comprising oil returns and EPU or EMU. With our data and sample, we obtain \( q = 9 \) in the case of EPU, whereas \( q = 5 \) for EMU. Differently to Balcilar et al. (2015), where the test is applied on the full sample, our analysis is carried out through the rolling window procedure above described. Consequently, on the one side, a large \( q \) would imply huge computational costs and, on the other side, we have, most likely, that \( q \) would change from one window to another. For that reason, as rule of thumb, and again to make a balance between precision of the analyses and computational burden, we set \( q = 2 \) in applying the causality test in quantiles. Likewise, for the models defined in (12)-(14), we set \( p = q = r = 2 \).

4 Empirical results

First of all, we analyze the causality in quantiles. In Figure 1 we report the values of the test statistic \( \hat{J}_T \), defined in (3), in case we study the causality implications of \( x_{1,t} \) on \( y_t \); differently, Figure 2 displays the output of the test applied for \( x_{2,t} \). The results in Figures 1-2 are very similar: periods in which \( \hat{J}_T \) takes low values (pointing out the low or inexistent power of the two uncertainty indices in causing the oil returns) are followed by periods of relevant peaks, such as in the second half of the 1980s, at the beginning and at the end of the 1990s, between the years 2006-2008. Moreover, we can see that the causality relations are stronger at the central \( \tau \) levels. Despite the regimes change over time, the periods in which the uncertainty indices are significant in causing the oil returns are less persistent than the ones characterized by no causality.

In Figures 3-4 we study the causality relations of \( x_{1,t} \) and \( x_{2,t} \), respectively, on \( y_t^2 \). Here, we observe a stronger causality impact with respect to the case of \( y_t \), since the periods in which \( x_{1,t} \) and \( x_{2,t} \) are significant in causing \( y_t^2 \) are more persistent. Once again, the causality relations are stronger at the central levels of \( \tau \).

We now define the dummy variable \( D_{y_t \cdot x_{1,t}}(\tau) \), taking value 1 if the test statistic \( \hat{J}_T \), defined in (3) and applied on the pair \((x_{1,t}, y_t)\), at the \( \tau \) level, as discussed above, is greater than 1.96, 0 otherwise. Likewise, \( D_{y_t \cdot x_{2,t}}(\tau) \), \( D_{y_t^2 \cdot x_{1,t}}(\tau) \) and \( D_{y_t^2 \cdot x_{2,t}}(\tau) \) are computed by using the same methodology and are obtained from the pairs \((x_{2,t}, y_t)\), \((x_{1,t}, y_t^2)\) and \((x_{2,t}, y_t^2)\), respectively. Figure 5 reports the linear correlation coefficients of those dummy variables, denoted as \( \rho_{D_{y_t}}(\tau) = \rho(D_{y_t \cdot x_{1,t}}(\tau), D_{y_t \cdot x_{2,t}}(\tau)) \) and \( \rho_{D_{y_t^2}}(\tau) = \rho(D_{y_t^2 \cdot x_{1,t}}(\tau), D_{y_t^2 \cdot x_{2,t}}(\tau)) \); we can see that the correlations, computed at different \( \tau \) levels, are not negligible, mainly in the case of \( y_t^2 \).

It is also important to evaluate the correlations among the variables of interest. In Figure 6 we show the trend of the linear correlation coefficients, computed through a rolling window procedure, with window size of 500 observations and step of one period ahead.
Figure 1: The causality of the Economic Policy Uncertainty (in logarithm) on the oil returns quantiles. The figure reports the values of the test statistic (3), computed through a rolling procedure with window size of 500 observations and step of 5 periods ahead.

Figure 2: The causality of the Market Equity Uncertainty (in logarithm) on the oil returns quantiles. The figure reports the values of the test statistic (3), computed through a rolling procedure with window size of 500 observations and step of 5 periods ahead.

It is important to see that all the correlations are not constant over time. $\rho(x_{1,t}, x_{2,t})$ always records the highest values; furthermore, $x_{1,t}$ and $x_{2,t}$ are more correlated with $y_{t}^2$ than $y_{t}$. Moreover, given the correlation coefficients obtained through the rolling window procedure, we compute their conditional average values. In particular, the correlation coefficients are computed by conditioning each pair of variables on the values of $y_t$ and $y_t^2$, respectively, such that $y_t$ and $y_t^2$ are lower or greater than their
The causality of Economic Policy Uncertainty (in logarithm) on the squared oil returns quantiles. The figure reports the values of the test statistic (3), computed through a rolling procedure with window size of 500 observations and step of 5 periods ahead.

The causality of the Market Equity Uncertainty (in logarithm) on the squared oil returns quantiles. The figure reports the values of the test statistic (3), computed through a rolling procedure with window size of 500 observations and step of 5 periods ahead.

The $\tau$-th in-sample quantiles, for $\tau = \{0.1, 0.5, 0.9\}$. The results are given in Table 4. We can see that, for each pair of variables and for each $\tau$, the results deeply differ depending on whether we condition the
Figure 5: The figure reports the linear correlation coefficients \( \rho_{Dy_t}(\tau) = \rho(Dy_t,x_1,t(\tau),Dy_t,x_2,t(\tau)) \) and \( \rho_{Dy_2^2}(\tau) = \rho(Dy_2^2(\tau),x_1,t,Dy_2^2(\tau),x_2,t(\tau)) \). \( Dy_t,x_1,t(\tau) \) is a dummy variable taking value 1 if the test statistic (3), applied for the pair \((x_1,t,y_t)\) at the \( \tau \) level, and computed with window size of 500 observations and step of 5 periods ahead, is greater than 1.96, 0 otherwise. \( Dy_t,x_2,t(\tau),Dy_2^2,x_1,t(\tau) \) and \( Dy_2^2,x_2,t(\tau) \) are computed in the same way.

Figure 6: The dynamic correlations, \( \rho(\cdot) \), of the variables \( y_t, y_t^2, x_{1,t}, x_{2,t} \). The linear correlations coefficients are computed through a rolling window procedure with window size of 500 observations and step of one period ahead.

correlations coefficients on the values of \( y_t \) or \( y_t^2 \) being greater or lower than their respective in-sample \( \tau \) quantiles.

The results discussed above highlight the importance of considering the joint impact of EPU and EMU in forecasting the oil movements, as well as the need of using the quantile regression method, given the asymmetric relations among the variables at different \( \tau \) levels.
It is possible to observe a precise trend of the coefficients values over time, the explanatory variables are not always statistically significant in explaining the conditional quantiles of the response variable. After estimating the parameters of Models (12)-(14) for each of the subsamples determined by the rolling window procedure, we computed their respective average values and standard deviations. We checked that all the coefficients’ p-values, on average, are greater than 0.05, pointing out that, over time, the explanatory variables are not always statistically significant in explaining the conditional quantiles of $y_t$ and $y_t^2$. For that reason, we report in Table 5 the mean (columns 2-7) and the standard deviation (columns 8-13) of the coefficients conditional on the fact that their respective p-values are less or equal than 0.05; those average values are denoted by $\bar{\beta}_j(t)$, $\bar{\delta}_j(t)$, $\bar{\lambda}_j(t)$, whereas the standard deviations are denoted as $\sigma_{\beta_j(t)}$, $\sigma_{\delta_j(t)}$, $\sigma_{\lambda_j(t)}$, for $j = \{1, 2\}$. For simplicity, we display the results obtained at $\tau = \{0.1, 0.5, 0.9\}$.

The table reports the average correlations among the variables $y_t$, $y_t^2$, $x_{1,t}$ and $x_{2,t}$. The correlation coefficients are computed by conditioning the pairs of the variables on the values of $y_t$ and $y_t^2$, respectively, such that $y_t$ and $y_t^2$ are lower or greater than their $\tau$-th quantiles, for $\tau = \{0.1, 0.5, 0.9\}$.

Table 4: Conditional average correlations.

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>lower</th>
<th>greater</th>
<th>lower</th>
<th>greater</th>
<th>lower</th>
<th>greater</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho(y_t, x_{1,t})$</td>
<td>-0.1415</td>
<td>0.0286</td>
<td>-0.0837</td>
<td>0.0785</td>
<td>-0.0474</td>
<td>0.1428</td>
</tr>
<tr>
<td>$\rho(y_t, x_{2,t})$</td>
<td>-0.1750</td>
<td>0.0272</td>
<td>-0.1023</td>
<td>0.0941</td>
<td>-0.0716</td>
<td>0.1535</td>
</tr>
<tr>
<td>$\rho(y_t^2, x_{1,t})$</td>
<td>-0.0600</td>
<td>0.0610</td>
<td>0.0064</td>
<td>0.0853</td>
<td>0.0179</td>
<td>0.1277</td>
</tr>
<tr>
<td>$\rho(y_t^2, x_{2,t})$</td>
<td>-0.0649</td>
<td>0.0810</td>
<td>-0.0176</td>
<td>0.1140</td>
<td>0.0190</td>
<td>0.1993</td>
</tr>
</tbody>
</table>

The table reports the average values (%), in columns 2-7, and the standard deviations (%), in columns 8-13, computed by conditioning the pairs of the variables on the values of $y_t$ and $y_t^2$. The correlation coefficients are computed by conditioning the pairs of the variables on the values of $y_t$ and $y_t^2$, respectively, such that $y_t$ and $y_t^2$ are lower or greater than their $\tau$-th quantiles, for $\tau = \{0.1, 0.5, 0.9\}$.

After estimating the parameters of Models (12)-(14) for each of the subsamples determined by the rolling window procedure, we computed their respective average values and standard deviations. We checked that all the coefficients’ p-values, on average, are greater than 0.05, pointing out that, over time, the explanatory variables are not always statistically significant in explaining the conditional quantiles of $y_t$ and $y_t^2$. For that reason, we report in Table 5 the mean (columns 2-7) and the standard deviation (columns 8-13) of the coefficients conditional on the fact that their respective p-values are less or equal than 0.05; those average values are denoted by $\bar{\beta}_j(t)$, $\bar{\delta}_j(t)$, $\bar{\lambda}_j(t)$, whereas the standard deviations are denoted as $\sigma_{\beta_j(t)}$, $\sigma_{\delta_j(t)}$, $\sigma_{\lambda_j(t)}$, for $j = \{1, 2\}$. For simplicity, we display the results obtained at $\tau = \{0.1, 0.5, 0.9\}$.

Table 5: Quantile regression output.

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>$\bar{\beta}_1(t)$</th>
<th>$\bar{\beta}_2(t)$</th>
<th>$\bar{\delta}_1(t)$</th>
<th>$\bar{\delta}_2(t)$</th>
<th>$\bar{\lambda}_1(t)$</th>
<th>$\bar{\lambda}_2(t)$</th>
<th>$\sigma_{\beta_1(t)}$</th>
<th>$\sigma_{\beta_2(t)}$</th>
<th>$\sigma_{\delta_1(t)}$</th>
<th>$\sigma_{\delta_2(t)}$</th>
<th>$\sigma_{\lambda_1(t)}$</th>
<th>$\sigma_{\lambda_2(t)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.024</td>
<td>17.893</td>
<td>-0.835</td>
<td>-0.040</td>
<td>-0.529</td>
<td>-0.020</td>
<td>19.122</td>
<td>5.997</td>
<td>0.399</td>
<td>0.694</td>
<td>0.207</td>
<td>0.683</td>
</tr>
<tr>
<td>0.5</td>
<td>-11.531</td>
<td>-7.864</td>
<td>-0.404</td>
<td>0.291</td>
<td>0.011</td>
<td>0.130</td>
<td>1.613</td>
<td>0.543</td>
<td>0.139</td>
<td>0.332</td>
<td>0.257</td>
<td>0.260</td>
</tr>
<tr>
<td>0.9</td>
<td>-16.625</td>
<td>-18.386</td>
<td>0.550</td>
<td>0.829</td>
<td>0.555</td>
<td>0.765</td>
<td>15.213</td>
<td>6.581</td>
<td>0.146</td>
<td>0.414</td>
<td>0.401</td>
<td>0.119</td>
</tr>
</tbody>
</table>

The table reports the average values (%), in columns 2-7, and the standard deviations (%), in columns 8-13, computed by conditioning the pairs of the variables on the values of $y_t$ and $y_t^2$. The correlation coefficients are computed by conditioning the pairs of the variables on the values of $y_t$ and $y_t^2$, respectively, such that $y_t$ and $y_t^2$ are lower or greater than their $\tau$-th quantiles, for $\tau = \{0.1, 0.5, 0.9\}$.

Starting with the estimation of Model (12), on average, the impact of the explanatory variables changes according to the $\tau$ levels, an evidence against the so-called location-shift hypothesis, which assumes homogeneous effects of the covariates across the conditional quantiles of the response variable. It is possible to observe a precise trend of the coefficients values over $\tau$: negative for $\bar{\beta}_j(t)$, positive for $\bar{\delta}_j(t)$, $\bar{\lambda}_j(t)$, $j = \{1, 2\}$. On average, the lags of $y_t$ have a positive impact on the left tail of the response variable conditional distribution; on the other hand, their effects become negative at medium-high $\tau$ levels. This is expected as past negative returns lead to the increase of the series dispersion and thus
move the 0.1 (0.9) quantile further on the left (right), with an additional effect on the median. On the contrary, positive returns shrink the density toward the median, which is also moving to the right. We interpret these evidences as a form of asymmetry, where the sign of the shocks lead to opposite effects on the quantiles, and thus on the distribution, of the target variable.

The opposite phenomenon is observed for $x_{1,t-j}$ and $x_{2,t-j}$, $j = \{1, 2\}$; for the uncertainty indexes, we were expecting those signs. In fact, an increase in the uncertainty, moves the lower quantiles to the left and the upper quantiles to the right, with the impact on the median being smaller than that on other quantiles for $j = 1$.

With the exception of $x_{2,t-j}$, $j = \{1, 2\}$, the coefficients of the other explanatory variables are less volatile at the central levels of $\tau$. In Table 6 we report the number of subsamples in which each coefficient turns out to be statistically significant at the level of 0.05. It is possible to see that, at $\tau$ equal to 0.1, 0.5 and 0.9, $x_{2,t-2}$, $x_{2,t-1}$ and $y_{t-1}$ record, respectively, the highest number of periods in which their coefficients are statistically significant.

Moving to the estimation of Model (14), just $\delta_2(0.1)$, $\lambda_1(0.1)$ and $\lambda_2(0.1)$ are negative; nevertheless those coefficients take very low values. With the exception of $\lambda_2(\tau)$, all the other coefficients exhibit, on average, an increasing trend over $\tau$. At $\tau$ equal to 0.1, 0.5 and 0.9, $y_{t-2}$, $x_{1,t-1}$ and $y_{t-1}$, record, respectively, the highest number of periods in which their coefficients are statistically significant at the 5% level, as it is possible to see from Table 6.

### Table 6: Persistence of significance over the rolled windows.

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>$n_{\beta_1}(\tau)$</th>
<th>$n_{\delta_2}(\tau)$</th>
<th>$n_{\lambda_1}(\tau)$</th>
<th>$n_{\delta_2}(\tau)$</th>
<th>$n_{\lambda_1}(\tau)$</th>
<th>$n_{\lambda_2}(\tau)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>112</td>
<td>37</td>
<td>165</td>
<td>114</td>
<td>54</td>
<td>190</td>
</tr>
<tr>
<td>0.5</td>
<td>102</td>
<td>9</td>
<td>30</td>
<td>64</td>
<td>113</td>
<td>35</td>
</tr>
<tr>
<td>0.9</td>
<td>322</td>
<td>156</td>
<td>93</td>
<td>47</td>
<td>105</td>
<td>53</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Estimates of Model (14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
</tr>
<tr>
<td>0.5</td>
</tr>
<tr>
<td>0.9</td>
</tr>
</tbody>
</table>

The table reports the number of subsamples, determined by the rolling window procedure, in which each coefficient turns out to be statistically significant at the level of 0.05. The rolling window procedure is applied by using a window size of 500 observations and step of 5 days ahead.

The larger impact of the squared lagged returns on upper quantiles is again expected, signaling that large movements (either positive or negative) lead to a huge increase of the risk. We observe a similar patter on the uncertainty indexes, then increase has a large impact on the upper quantiles of the squared returns.

Subfigures 7(a)-7(c) show the conditional distributions of $y_t$ and $y_t^2$, respectively, estimated from the first of the rolled windows. We clearly note the problem of crossing in quantiles vanishes by applying the quantile bootstrap method proposed by Chernozhukov et al. (2010). Moreover, the Epanechnikov kernel method allows to obtain smoother distributions. Subfigures 7(b)-7(d) show the conditional
densities of \( y_t \) and \( y_t^2 \), respectively, estimated from the first and the last windows of the rolling window procedure, by applying the Epanechnikov kernel method. Notably, the shape of each density changes over time, thus supporting the need for a rolling evaluation.

Figure 7: Conditional distributions and densities of \( y_t \) and \( y_t^2 \). Subfigures (a) and (c) display, respectively, the conditional distributions of \( y_t \) and \( y_t^2 \), estimated from the first subsample determined through the rolling window procedure. “Original”, “Adjusted” and “Kernel” stand for the distributions arising directly from Models (12)-(14), the ones obtained by adjusting the original estimates through the quantile bootstrap method proposed by Chernozhukov et al. (2010), and the ones built by means of the Epanechnikov kernel, respectively. Subfigures (b) and (d) show the conditional densities of \( y_t \) and \( y_t^2 \), respectively, estimated from the first and the last windows determined by the rolling window procedure, by applying the Epanechnikov kernel method.

We now evaluate the possible asymmetric effects of the uncertainty indices on the oil movements. In doing that, we slightly modify Models (12) and (14); first of all, we center to zero both \( x_{1,t-j} \) and \( x_{2,t-j} \), \( j = \{1, 2\} \), by subtracting from them their respective average values. Those new variables, which now can take both positive and negative values, are denoted by \( x^*_{1,t-j} \) and \( x^*_{2,t-j} \), \( j = \{1, 2\} \), respectively. Secondly, we make use of the following indicator functions: \( 1_{\{x^*_{1,t-j}<0\}} \) and \( 1_{\{x^*_{2,t-j}<0\}} \), \( j = \{1, 2\} \), which take value 1 if the condition in \{·\} is true, 0 otherwise. Specifically, the new models are defined as follows:

\[
Q_\tau(y_t|W_{t-1}) = \alpha_0(\tau) + \beta_1(\tau)y_{t-1} + \beta_2(\tau)y_{t-2} + \delta_1^d(\tau)x^*_{1,t-1} + \delta_2^d(\tau)x^*_{1,t-2} + \lambda_1^d(\tau)x^*_{2,t-1} \\
+ \lambda_2^d(\tau)x^*_{2,t-2} + \delta_1^1(\tau)1_{\{x^*_{1,t-1}<0\}}x^*_{1,t-1} + \delta_2^1(\tau)1_{\{x^*_{1,t-2}<0\}}x^*_{1,t-2} \\
+ \lambda_1^1(\tau)1_{\{x^*_{2,t-1}<0\}}x^*_{2,t-1} + \lambda_2^1(\tau)1_{\{x^*_{2,t-2}<0\}}x^*_{2,t-2},
\]  

(32)
In evaluating the asymmetric effects of EPU (EMU) on the oil movements, it is important to notice that the impact of \( x_{1,t-j} \) (\( x_{2,t-j} \)), \( j = \{1, 2\} \), is quantified by \( \delta_d^j(\tau) (\lambda_d^j(\tau)) \) if \( x_{1,t-j} \geq 0 (x_{2,t-j} \geq 0) \); differently, its impact is equal to \( \delta_d^j(\tau) + \delta_u^j(\tau) (\lambda_d^j(\tau) + \lambda_u^j(\tau)) \) if \( x_{1,t-j} < 0 (x_{2,t-j} < 0) \).

We report the results arising from the estimation of Models (32)-(33) in Table 7. Here, we display the average values of the coefficients over the rolled subsamples (window size of 500 observations and steps of 5 days ahead), conditioned to the fact that they are statistically significant at the level of 5%; we also report their standard deviations. More precisely, for instance, in order to evaluate correctly the asymmetric effects of \( x_{1,t-1} \), for each window, we considered the cases where all the coefficients \( \left[ \hat{\delta}_1^j(\tau), \hat{\delta}_u^j(\tau) = \hat{\delta}_d^j(\tau) + \hat{\delta}_u^j(\tau) \right] \) are simultaneously significant, and then we computed their average values. Likewise, we applied the same methodology for the other coefficients estimated from Models (32)-(33).

### Table 7: The asymmetric impact of uncertainty on the oil movements.

<table>
<thead>
<tr>
<th>( \tau )</th>
<th>( \hat{\delta}_d^j(\tau) )</th>
<th>( \hat{\delta}_u^j(\tau) )</th>
<th>( \hat{\lambda}_d^j(\tau) )</th>
<th>( \hat{\lambda}_u^j(\tau) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimates of Model (32)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>-2.44 (1.9)</td>
<td>-0.12 (0.4)</td>
<td>-1.16 (0.6)</td>
<td>-1.07 (1.0)</td>
</tr>
<tr>
<td>0.5</td>
<td>0.95 (0.4)</td>
<td>-0.35 (0.2)</td>
<td>-0.38 (0.2)</td>
<td>-0.92 (0.8)</td>
</tr>
<tr>
<td>0.9</td>
<td>2.20 (1.9)</td>
<td>-0.92 (0.8)</td>
<td>1.45 (1.1)</td>
<td>-0.65 (1.9)</td>
</tr>
<tr>
<td>Estimates of Model (33)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>1.00 (8.2)</td>
<td>0.08 (8.2)</td>
<td>-1.02 (2.2)</td>
<td>-2.92 (10.2)</td>
</tr>
<tr>
<td>0.5</td>
<td>2.68 (0.8)</td>
<td>-0.30 (1.4)</td>
<td>-2.07 (2.6)</td>
<td>-0.17 (0.3)</td>
</tr>
<tr>
<td>0.9</td>
<td>14.14 (12.0)</td>
<td>-0.52 (3.1)</td>
<td>29.51 (15.2)</td>
<td>-2.90 (3.9)</td>
</tr>
</tbody>
</table>

The table reports the average values, computed over the subsamples determined by the rolling window procedure, of the uncertainty indices’ coefficients (in brackets we report their standard deviations) estimated for Models (32) and (33), conditional to the fact that they are statistically significant at the level of 0.05. The rolling window procedure is applied by using a window size of 500 observations and step of 5 days ahead. \( \hat{\delta}_u^j(\tau) \) is the conditional average value of the sums \( \hat{\delta}_d^j(\tau) + \hat{\delta}_u^j(\tau) \), computed from the windows where the two coefficients are simultaneously significant; similarly, \( \hat{\lambda}_u^j(\tau) \) is the conditional average value \( \hat{\lambda}_d^j(\tau) + \hat{\lambda}_u^j(\tau) \), \( j = \{1, 2\} \).

From Table 7, it is possible to see that, on average, both the uncertainty indices have asymmetric effects on the oil movements, and that the impact is stronger in the states where they take high values, i.e. when \( x_{1,t-j} \) and \( x_{2,t-j} \) take positive values. Indeed, the means of \( \hat{\delta}_d^j(\tau) \) and \( \hat{\lambda}_d^j(\tau) \) are almost always greater, in absolute value, than the means of \( \hat{\delta}_d^j(\tau) + \hat{\delta}_u^j(\tau) \) and \( \hat{\lambda}_d^j(\tau) + \hat{\lambda}_u^j(\tau) \), \( j = \{1, 2\} \), respectively. This is a somewhat expected result suggesting that increases in uncertainty do have a larger impact on oil movements compared to decreases in uncertainty.

The last point concerns the evaluation of the Models (12)-(14) predictive power, placing particular emphasis on the contribution of the two uncertainty indices \( x_{1,t} \) and \( x_{2,t} \) in forecasting the \( y_t \) and \( y_{1,t}^2 \) quantiles and distributions. The results arising from the Berkowitz (2001) test are given in Subfigures...
8(a), which displays the case of \( y_t \), and 8(b), where we focus on \( y^2_t \).

In both the cases we display the p-values obtained from 4 different models: Model 1 includes all the available predictors, i.e. \( y_{t-j}, x_{1,t-j}, x_{2,t-j} \) in Subfigure 8(a), \( y^2_{t-j}, x_{1,t-j}, x_{2,t-j} \) in Subfigure 8(b); Model 2 has just \( y_{t-j} \) in Subfigure 8(a), \( y^2_{t-j} \) in Subfigure 8(b); Model 3 comprises \( y_{t-j}, x_{1,t-j} \) in Subfigure 8(a), \( y^2_{t-j}, x_{1,t-j} \) in Subfigure 8(b); finally, Model 4 includes \( y_{t-j}, x_{2,t-j} \) in Subfigure 8(a), \( y^2_{t-j}, x_{2,t-j} \) in Subfigure 8(b). In all the cases we set \( j = \{1, 2\} \).

Figure 8: The p-values of the Berkowitz (2001) test over the rolled windows. Subfigure (a) and (b) refer, respectively, to the forecasts of the \( y_t \) and \( y^2_t \) distributions. Model 1 includes all the predictors in (12) and (14), respectively. Model 2 includes just \( y_{t-1} \) and \( y_{t-2} \), in “OIL RETURNS”, \( y^2_{t-1} \) and \( y^2_{t-2} \), in “SQUARED OIL RETURNS”. Model 3 includes \( y_{t-1}, y_{t-2}, x_{1,t-1} \) and \( x_{1,t-2} \), in “OIL RETURNS”, \( y^2_{t-1}, y^2_{t-2}, x_{1,t-1} \) and \( x_{1,t-2} \) in “SQUARED OIL RETURNS”. Finally, Model 4 includes \( y_{t-1}, y_{t-2}, x_{2,t-1} \) and \( x_{2,t-2} \), in “OIL RETURNS”, \( y^2_{t-1}, y^2_{t-2}, x_{2,t-1} \) and \( x_{2,t-2} \) in “SQUARED OIL RETURNS”.

For all the 4 models, there are periods in which the null hypothesis of correct specification is not rejected, and others where the null hypothesis is rejected. In some periods, the inclusion of the uncertainty indices implies evident benefits: here, the p-values generated by Models 1-3-4 are greater than 0.05, whereas those of Model 2 are less than 0.05. We can also observe that during the years 2003-2007 and 2009 (in Subfigure 8(a)), 1990-1991, 2004, 2009-2010 and 2014 (in Subfigure 8(b)), the null hypothesis is not rejected just for Model 1, highlighting the importance of exploiting the joint predictive power of EPU and EMU.

The direct comparisons between the restricted (the one including just the lags of \( y_t \)) and the unrestricted (which includes also the lags of \( x_{1,t} \) and \( x_{2,t} \)) models, based on the tests proposed by Diebold and Mariano (2002), Amisano and Giacomini (2007), Diks et al. (2011) and Gneiting and Ranjan (2011), are evaluated by means of Figures 9, 10, 11 and 12, respectively. As for the forecasting of the \( y_t \) quantiles and distribution (Subfigures 9(a), 10(a), 11(a) 12(a)), it is possible to see that the test statistics change their sign over time, pointing out periods in which the unrestricted model works better, providing the highest scores, followed by others where the best performance is recorded by the restricted model. Nevertheless, the null hypothesis of equal performance is not always rejected at the level of 5% and the periods in which the unrestricted model records the best performance, statistically significant, are less frequent than the ones where it is outperformed by the restricted model.
Figure 9: The Diebold and Mariano (2002) test statistic values over the rolled windows. Subfigure (a) and (b) refer, respectively, to the forecasts of the $y_t$ and $y_t^2$ distributions, where we compare the restricted (which contain just the lags of $y_t$ and $y_t^2$) and the unrestricted models (which include also the lags of $x_{1,t}$ and $x_{2,t}$). The test is applied at three different $\tau$ values: 0.05 (black lines), 0.50 (red lines) and 0.95 (blue lines).

Figure 10: The Amisano and Giacomini (2007) test statistic values over the rolled windows. Subfigure (a) and (b) refer, respectively, to the forecasts of the $y_t$ and $y_t^2$ distributions, where we compare the restricted (which contain just the lags of $y_t$ and $y_t^2$) and the unrestricted models (which include also the lags of $x_{1,t}$ and $x_{2,t}$). The test is applied by placing greater emphasis on the center (red lines), on the right tail (blue lines) and on the left tail (green lines) of the conditional distributions. It is also applied by placing equal weight to the different regions of the distributions (black lines).

Figure 11: The Diks et al. (2011) test statistic values over the rolled windows. Subfigure (a) and (b) refer, respectively, to the forecasts of the $y_t$ and $y_t^2$ distributions, where we compare the restricted (which contain just the lags of $y_t$ and $y_t^2$) and the unrestricted models (which include also the lags of $x_{1,t}$ and $x_{2,t}$). The test is applied by placing greater emphasis on the center (red lines), on the right tail (blue lines) and on the left tail (green lines) of the conditional distributions.
In general, all the tests give evidence of the best performance of the unrestricted model in the second half of the 2000s and at the end of the 2000s, when we place emphasis on the right tail of the $y_t$ conditional distribution. The tests proposed by Amisano and Giacomini (2007) and Diks et al. (2011) detect further periods, namely in the middle of the 1990s and at the beginning of the 2000s, where the unrestricted model outperforms the restricted one, mainly when we focus on the center of the distribution.

Now we collect the information coming from the five tests to highlight the periods where EPU and EMU turn out to be crucial in forecasting the $y_t$ conditional distribution and quantiles. For this purpose, we computed, for each of the applied tests (i.e. those developed by Berkowitz (2001), Diebold and Mariano (2002), Amisano and Giacomini (2007), Diks et al. (2011) and Gneiting and Ranjan (2011)) a dummy variable, $D_{t}^{pred}$, taking value 1 if the unrestricted model records a (statistically) better performance, at the level of 0.05, than the restricted one at $t$, 0 otherwise. In case of the Berkowitz (2001) test, $D_{t}^{pred}$ takes value 1 if the null hypothesis is not rejected for the unrestricted model (which includes all the available covariates) and rejected for the restricted one (which includes just the lagged values of $y_t$ or $y_t^2$) at $t$. In order to clean the series from the periods where the better performance of one of the two models lasts for a few day, being negligible, we compute, for each test, the following moving average:

$$SD_{t}^{pred} = \frac{1}{2M_s + 1}\sum_{t-M_s}^{t+M_s} D_{t}^{pred}.$$ \hspace{1cm} (34)

In our work we set $M_s = 10$, hence the moving averages span 21 days, and identify the periods where the unrestricted model provides the best performance as those where $SD_{t}^{pred} \geq 0.5$. In Figure 13 we display those periods with different colours for each test, linking them with the $y_t$, $x_{1,t}$ and $x_{2,t}$ series. It is possible to see that there exists a relevant evidence of a crucial role of EPU and EMU in forecasting the $y_t$ conditional quantiles and distributions, as highlighted at least by 3 of the implemented tests, during the years 2005-2007 and 2008-2010. Those periods are close to two special

Figure 12: The Gneiting and Ranjan (2011) test statistic values over the the rolled windows. Subfigure (a) and (b) refer, respectively, to the forecasts of the $y_t$ and $y_t^2$ distributions, where we compare the restricted (which contain just the lags of $y_t$ and $y_t^2$) and the unrestricted models (which include also the lags of $x_{1,t}$ and $x_{2,t}$). The test is applied by placing greater emphasis on the center (red lines), on the right tail (blue lines) and on the left tail (green lines) of the conditional distributions.

Figure 13: Evidence of the crucial impact of EPU and EMU in forecasting the $y_t$ conditional distribution and quantiles over time, detected by several tests and linked to the $y_t$ (a), $x_{1,t}$ (b), and $x_{2,t}$ (c) series. A, B, C, D, E, F, G, H, I stand for the tests proposed by Diks et al. (2011), with focus on the right and the center parts of the distribution, Gneiting and Ranjan (2011), with focus on the right tail of the distribution, Amisano and Giacomini (2007), with focus on the right, the center and the left parts of the distribution, Diebold and Mariano (2002), with focus on the 95-th and the 5-th percentile, Berkowitz (2001), respectively.

As regards the case of $y^2_t$, (Subfigures 9(b), 10(b), 11(b) 12(b)), the tests proposed by Diebold and Mariano (2002) and Gneiting and Ranjan (2011) lead to conclusions similar to those obtained for $y_t$; indeed there is evidence of a better performance of the unrestricted model in the years 2009-2010, when the emphasis is placed on the right tail. The Diebold and Mariano (2002)’s test detects such an evidence also at the beginning of the 1990s. Differently, according to the the tests developed by Amisano and Giacomini (2007) and Diks et al. (2011), the unrestricted model works better, with respect to the $y_t$ case, when we focus on the center and on the left tail of the $y^2_t$ distribution. Furthermore, on the basis of the Diks et al. (2011)’s test, the periods characterized by the best performance, statistically significant, of the unrestricted model, become more persistent with respect to the $y_t$ case.

Like the $y_t$ case, we collect the information coming from the several implemented tests and highlight the periods of the enduring best performance of the unrestricted model, by computing the quantity in (34). We display the results in Figure 14. Similarly to $y_t$, we have evidence of a crucial role of EPU and EMU in forecasting the $y^2_t$ conditional quantiles and distributions during the years 2005-2007 and 2008-2010.

To summarize, we checked that the relationships among the variables included in Models (12) and (14) change over time. As a result, the logarithms of EPU and EMU turn out to be relevant variables,
Evidence of the crucial impact of EPU and EMU in forecasting the \( y_t^2 \) conditional distribution and quantiles over time, detected by several tests and linked to the \( y_t \) (a), \( x_{1,t} \) (b), and \( x_{2,t} \) (c) series. A, B, C, D, E, F, G stand for the tests proposed by Diks et al. (2011), with focus on the center and the left parts of the distribution, Gneiting and Ranjan (2011), with focus on the right tail of the distribution, Amisano and Giacomini (2007), with focus on the center and the left parts of the distribution, Diebold and Mariano (2002), with focus on the 95-th percentile, Berkowitz (2001), respectively.

In causing and forecasting the conditional quantiles and distributions of \( y_t \) and \( y_t^2 \), just in some periods. Such an evidence supports the use of the rolling window procedure to capture those dynamics, instead of carrying out a full sample analysis. The periods in which the two uncertainty indices are significant in causing and in forecasting \( y_t^2 \), at different regions of its distribution, are more persistent with respect to the ones recorded in the case of \( y_t \); as for the forecasting exercise, this phenomenon is observed through the tests proposed by Amisano and Giacomini (2007) and Diks et al. (2011). The reason might be the following: with the squared returns of oil, we focus on the \( y_t \) volatility, a measure of dispersion (and thus uncertainty) which fits better the nature of EPU and EMU, which are themselves uncertainty indicators.

5 Concluding remarks

In the present work we checked that the relations among the oil movements, defined as returns (\( y_t \)) and squared returns (\( y_t^2 \)), and the uncertainty indices (EPU and EMU) are affected by structural breaks. The conclusions drawn from a full sample analysis would be misleading and, therefore, we implemented a rolling window procedure in order to capture the dynamics among the involved variables.

We first showed that the impact of EPU and EMU in causing the quantiles of the oil returns changes over time. Indeed, periods characterized by low or inexistent power of the two uncertainty
indices in causing the oil returns are followed by periods of relevant causality evidence. Nevertheless, despite the changing in regimes over time, the periods in which the uncertainty indices are significant in causing the oil returns are less persistent than the ones characterized by no causality. Differently, when we focused on the $y_t^2$ case, then considering a quantile causality in variance, we observed stronger causality impacts, since the periods in which $x_{1,t}$ and $x_{2,t}$ are significant in causing the $y_t^2$ quantiles are more persistent. In both the cases, the causality relations are stronger at central quantiles levels.

Similarly, EPU and EMU turned out to be important drivers in forecasting the $y_t$ and the $y_t^2$ conditional distributions just in some periods. Indeed, their coefficients are not always statistically significant, at the 5% level; as a result, the predictive power of the model we propose, evaluated ex post, from the out-of-sample realizations, is significantly improved by EPU and EMU just in some periods, as showed by several tests, namely those introduced by Berkowitz (2001), Diebold and Mariano (2002), Amisano and Giacomini (2007), Diks et al. (2011), Gneiting and Ranjan (2011). Moreover, those tests reveal different impacts according to the different regions of the response variables conditional distributions. In particular, through the Amisano and Giacomini (2007) and the Diks et al. (2011) testing approaches, we checked that, consistently to the causality analysis, the periods in which the two uncertainty indices are significant in forecasting $y_t^2$, at different regions of its distribution, are more persistent with respect to the ones recorded in the case of $y_t$. The reason might be the following: with the squared returns of oil, we focus on the $y_t$ volatility, a measure of dispersion (and thus uncertainty) which fits better the nature of EPU and EMU, which are themselves uncertainty indicators.

References


Balcilar, M., Bekiros, S., and Gupta, R. (2015). The role of news-based uncertainty indices in pre-


33