Speculative behaviour and oil price predictability

Ekaterini Panopoulou*

University of Piraeus, Greece

Theologos Pantelidis†

University of Macedonia, Greece

Abstract

We develop two- and three-state regime switching models and test their forecasting ability for oil prices. We use the deviations of market oil price from fundamental values as the main explanatory variable in our models, while additional potential predictors enrich our specification. Our findings suggest that the regime-switching models are, in general, more accurate than the Random Walk model in terms of both statistical and economic evaluation criteria for oil price forecasts.

JEL Classification: C5; Q4;
Keywords: Oil price; Regime Switching; Forecasting; Deviations from fundamentals.

* *Correspondence to:* Ekaterini Panopoulou, Department of Statistics and Insurance Science, University of Piraeus, 80 Karaoli & Dimitriou Str., 18534, Piraeus, Greece. E-mail: apano@unipi.gr. Tel: 0030 210 4142728. Fax: 0030 210 4142340.

† Theologos Pantelidis, Department of Economics, University of Macedonia, 156 Egnatia Str., GR-54006, Thessaloniki, Greece. E-mail: pantelidis@uom.gr. Tel.: 0030 2310 891726. Fax: 0030 2310 891292.

1 Introduction

Oil price is a key-variable in macroeconomic projections affecting inflation and economic activity. Clearly, the predictability of the price of oil is of great interest to policymakers, central banks, CEOs and international investors. Strategic and investment decisions of airline, automobile and energy companies are based on scenarios built on forecasts of the future path of oil price. Even homeowners have in mind some kind of expectations about the future price of oil when deciding about energy-saving investments. Moreover, energy and especially crude oil futures have become widespread investment vehicles among traditional and alternative asset managers, mainly due to their equity-like return, their inflation-hedging properties and their role for risk diversification.

The recent surge in oil prices (and other commodities as well) between 2003-2008 has sparked a public debate on the determinants of the price of crude oil. Fundamental-based explanations of oil price movements are attributed to oil supply shocks, oil demand shocks driven by global economic activity, and oil-specific demand shocks. Oil supply shocks stem from reduced oil production of oil-exporting regions, while an oil demand shock is mainly caused by unexpected world economic activity. Finally, an oil-specific demand shock may be triggered by either changing expectations of oil fundamentals or financial speculation. It seems that the literature has reached a consensus on the drivers of the oil price boom until mid-2008. Specifically, Hamilton (2009 a,b), Kilian, (2009), Kilian and Hicks (2009), Juvenal and Petrella (2012) and Kilian and Murphy (2013) find that the recent oil price rise is mainly attributed to strong oil demand confronting stagnating global oil production. With respect to the oil-specific demand shock, this can be decomposed into an oil-specific shock which captures changes in oil demand not related to economic activity, and a destabilising financial shock. Lombardi and Van Robays (2011) attempt such a decomposition and model the destabilising financial shock as a shock that creates a perturbation in the futures market due to increased demand for futures contracts that moves the futures price away from its efficient level. Such financial shocks may emerge due to the increasing financialisation of oil futures markets measured by the sharp rise in speculative open interest and speculative market shares (see among others Mayer, 2010; Irwin and Sanders, 2011; Tang and Xiong, 2011; CFTC, 2011 Fattouh *et al.* 2013). However, it is not clear whether the way market participants act is due to lack of a fundamental basis in supply and demand or whether it represents the mechanism by which market fundamentals are incorporated in competitively determined prices. Kilian and Murphy (2013) and Kilian and Lee (2013) argue that financialisation in oil futures markets should be modelled as part of the endogenous propagation of shocks to fundamentals rather than an exogenous intervention.

Speculative behaviour, however, can cause oil price changes that generate bubbles. In the case of rational bubbles, these are generated by endogenous responses to the fundamentals that drive asset prices (Branch and Evans, 2011). The literature mainly focuses on speculative bubbles in the stock market, while the evidence for the oil market is quite scanty. We mention three related studies which explicitly test for speculative bubbles in the oil price by making use of the recently proposed Supremum Augmented Dickey Fuller (SADF) approach (proposed by Phillips *et al.*, 2011). Specifically, Gilbert (2010) and Homm and Breitung (2012) cannot detect speculative bubbles in oil prices consistently. By contrast, Phillips and Yu (2011) succeed in detecting explosive behaviour in monthly oil prices normalised by US inventories between March and July 2008. Applying the duration dependence test, Went *et al.* (2012) provide further evidence in favour of speculative bubbles in the oil price. Einloth (2009) also attributes part of the oil price movements in recent years to speculation. More recently, Lammerding *et al.* (2013) draw on the relationship between oil prices and oil dividends and establish a state-space framework from which they extract the bubble component as an unobservable variable. They additionally assume the bubble to evolve over time as a two-state Markov-switching process with two distinct regimes; namely one in which the bubble evolves over time as a stable process and one in which the bubble exhibits explosive dynamics. The authors follow a Bayesian approach, implementing a fully-fledged Markov-Chain-Monte-Carlo (MCMC) estimation framework and find convincing evidence of two distinct bubble episodes in the oil market.

In this study, we employ and develop models of speculative behaviour in the oil market building on the existence of a bubble. A word of caution is in order here. We should note that our models do not allow us to attribute the source of a bubble to specific characteristics of the oil market and as such we can not infer whether a bubble is based on fundemental or non-fundamental factors. In any event, we do not attempt to discriminate between the two hypotheses. We follow Pindyck (1993) and infer the fundamental value of crude oil from the current and expected discounted convenience yield that accrues from holding inventories based on a non-arbitrage condition between oil spot and futures price. Any deviation of current values from fundamental values is termed 'bubble' and may summarise a variety of shocks as outlined before. The bubble component can be in one of two or three regimes giving rise to our two- and three-state Regime-Switching (RS) models along the lines of the models developed by Van Norden and Schaller (1993, VNS hereafter) and Brooks and Katsaris (2005, BK hereafter). The authors link speculative behaviour in asset returns to RS models. Specifically, VNS show that a two-regime speculative behaviour model, in which the bubble is allowed to switch between a survival and a collapse state has significant explanatory power for stock returns. BK incorporate a third regime in the VNS model to allow for the bubble growing at a steady rate of return bridging the gap between VNS and Evans (1991) who allow for the bubble to switch between the dormant and the explosive state. Recently, Shi and Arora (2012, SA hereafter) extended the VNS and BK models to oil prices and found a reasonably good fit of the data along with evidence of a speculative bubble over the late-2008/early-2009 period.

The aforementioned studies focus on the in-sample ability of RS models to capture the dynamics in the price of the asset under scrutiny, ignoring the out-of-sample predictive power of the models. This is the goal of our analysis. To be more specific, we augment both the VNS and BK models by adding a variety of variables that serve as predictors of the future dynamics in the oil price and investigate the forecasting performance of these specifications. Following BK and SA, we employ the abnormal futures trading volume as a signal of market expectations governing both the mean and the probability equation of the surviving regime. This variable can be thought of as a destabilising financial shock in the context of Lombardi and Van Robays (2011). In a similar manner, we incorporate the variables proposed by Kilian and Murphy (2013) which are linked to oil supply shocks, oil demand shocks and oil-specific (speculative) demand shocks. Widening the information set (see also Juvenal and Petrella, 2012), we also employ macroeconomic and financial variables which act complementarily to measures of economic activity and financial conditions.

The forecasting performance of our models is evaluated in both statistical and economic terms. Economic evaluation is desirable since the oil market and the commodities markets have attracted the interest of large financial institutions, hedge funds and investment funds in general. Commodities are included in investment portfolios in order to diversify risk (Gorton and Rouwenhorst 2006). To anticipate our key results, the RS models appear to generate more accurate forecasts of the oil price, in both statistical and economic evaluation terms, relative to the Random Walk (RW) benchmark. Specifically, the RS models considered in this study outperform the RW model and the improvement in the accuracy of the oil price forecasts is statistically significant in all cases. Moreover, their superiority over the RW model is even stronger in economic evaluation terms. Finally, many of the predictors examined in this study appear to improve the forecasting accuracy of the RS models.

In the literature there are many studies that focus on oil forecasts but, to the best of our knowledge, none of them employs the class of RS models considered in our study. For example, Knetsch (2007) generates forecasts of the price of oil by means of a convenience yield forecasting model. His approach leads to more accurate forecasts of oil prices compared to direct forecasts from futures prices of the commodity but fails to beat the RW model (based on the root mean squared error criterion). Similarly, Alquist and Kilian (2010) provide evidence that forecasts from oil futures prices tend to be less accurate than forecasts from the RW model. Wu and McCallum (2005) argue that the accuracy of oil price forecasts can be improved by taking into account the relationship between current spot and futures prices instead of considering only the raw futures price. Baumeister and Kilian (2012) organise a forecasting exercise in real-time terms and provide evidence supporting the ability of Vector AutoRegressive (VAR) models to generate reliable forecasts of the real price of oil, while Baumeister and Kilian (2013) examine the predictability of the oil price from a central banker's point of view. Finally, Alquist *et al.* (2011) provide a stimulating review on the predictability of oil price.

The remainder of this paper is organised as follows. Section 2 introduces the Regime-Switching models used in this study, describes our approach to construct fundamental values and outlines the rationale behind the choice of predictors included in our models. Section 3 describes the dataset and reports the empirical findings of the study. Finally, Section 4 concludes.

2 Economic modelling and econometric specification

In this section, we initially provide a brief description of the three-state RS model of BK, which we augment with various predictors for the price of oil. The selection of these predictors is based on the existing literature on the determinants of oil price. We also describe two restricted versions of the three-regime model which we consider in our study and apply an arbitrage relation to calculate the convenience yield that allows us to obtain fundamental values.

2.1 Speculative behaviour and regime-switching models

Consider a simple asset pricing model where risk-neutral investors choose between holding an asset that yields $(1+r)$ and a risky asset, in our case oil. The investors' first order conditions imply that the price of the asset, P_t is given as follows:

$$P_t = \frac{1}{1+r}(E_t(P_{t+1} + D_t)$$

where D_t is some payoff in the form of dividends (stock market), convenience yield (oil and commodity markets), etc. One possible solution of the above equation defines the fundamental value of the asset, P_t^f, as follows:

$$P_t^f = \sum_{k=0}^{\infty}(1+r)^{-k+1}E_t(D_{t+k})$$

All other prices are said to be bubbly and the bubble component, (B_t), is defined as[1]

$$B_t = P_t - P_t^f$$

We assume that the expected size of the bubble in the next period can be generated from one of three regimes: a deterministic (dormant) regime (D), a surviving explosive regime (S) and a collapsing explosive regime (C) giving rise to a three regime model of speculative behaviour. As BK suggest, while several variables may prove significant in classifying the behaviour of the asset of interest, the relative size of the bubble is expected to play a predominant role.

If the bubble is in regime D, investors believe that it will continue to grow at a constant rate $(1 + r_f)$:

$$E_t(B_{t+1} \mid W_{t+1} = D) = (1 + r_f)B_t \tag{1}$$

where B_t is the size of the bubble (the difference between the actual and fundamental values) at time t, r_f is a constant discount rate and W_t (or $State_t$) is an unobserved variable that determines the regime. In this regime, the probability of collapse is negligible and investors do not demand an excess return for this. The probability of being in regime D in time $t + 1$ is denoted by η_t and depends on the relative size of the bubble (b_t) and other observed variables at time t. Even when the bubble is in the dormant regime, investors attach some probability in the bubble entering the explosive state by either surviving or collapsing. The probability of the explosive state is given by $1 - \eta_t$ and in this explosive state the probabilities of the two underlying regimes (Survive or Collapse) are denoted by q_t and $(1- q_t)$, respectively. In other words, the probability of being in each state is as follows:

$$\begin{aligned}
\Pr(State_{t+1} &= D) = \eta_t \\
\Pr(State_{t+1} &= S) = (1 - \eta_t)q_t \\
\Pr(State_{t+1} &= C) = (1 - \eta_t)(1 - q_t)
\end{aligned}$$

In the collapsing regime, the expected size of the bubble is given by

$$E_t(B_{t+1} \mid W_{t+1} = C) = g(b_t)P_t \tag{2}$$

where $g(b_t)$ is a continuous and everywhere differentiable function such that $g(0) = 0$ and $0 \preceq \partial g(b_t)/\partial b_t \preceq 1 + r_f$, b_t is the relative size of the bubble ($b_t = B_t/P_t$) and P_t is the

[1]The fundamental component, and consequently the bubble component, is calculated by means of the present value model of rational commodity pricing put forward by Pindyck (1993). Please refer to Subsection 2.2 for details on the calculation of the bubble.

actual asset price at time t.[2] Finally, the expected value of the bubble in the surviving regime is given by:[3]

$$E_t(B_{t+1} \mid W_{t+1} = S) = \frac{(1+r_f)}{q_t} B_t - \frac{(1-q_t)}{q_t} g(b_t) P_t \tag{3}$$

After some manipulation, the expected values in equations (1)-(3) can be written in terms of expected gross total returns of the next period, R_{t+1} as follows:

$$E_t(R_{t+1} \quad | \quad W_{t+1} = D) = (1+r_f) \tag{4}$$

$$E_t(R_{t+1} \quad | \quad W_{t+1} = C) = (1+r_f)(1-b_t) + g(b_t) \tag{5}$$

$$E_t(R_{t+1} \quad | \quad W_{t+1} = S) = (1+r_f) + \frac{(1-q_t)}{q_t}[(1+r_f)b_t - g(b_t)] \tag{6}$$

where $R_{t+1} = (P_{t+1} + D_{t+1})/P_t$. Equations (4)-(6) suggest that in the dormant regime investors expect a gross return equal to the required rate of return on the bubble-free asset, while in the surviving and collapsing regimes the gross return is a function of both the required rate of return and the relative size of the bubble.

We now turn to modelling the probabilities η_t and q_t which as already mentioned are negative functions of the bubble and specifically the absolute value of the bubble, $|b_t|$, since we allow for both negative and positive deviations. In order to model the probability of being in the dormant regime in the next time period, we follow BK and include the absolute value of the spread, (s_t) defined as the average 12-month actual returns minus the absolute value of the average 12-month returns of the estimated fundamental values. The intuition behind this specification is quite clear. When investors observe large spreads, i.e. larger average returns than average fundamental returns, they tend to believe that the bubble has entered the explosive state and the probability of being in the dormant state falls. In order to ensure that the estimated probability is bounded between 0 and 1, we adopt a Probit specification. Under this setting, the probability of being in the dormant regime in the next time period is given by:

$$\Pr(State_{t+1} = D) = \eta_t = \Phi(\beta_{\eta_0} + \beta_{\eta_1} |b_t| + \beta_{\eta_2} s_t) \tag{7}$$

where Φ is the cumulative density function of the standard normal distribution.

The basic assumption of the speculative bubble models is that the arrival of news may fuel a bubble collapse, which is often viewed as a random occasion that causes investors to liquidate their position at a certain point in time. Although investors observe the

[2]Please note that the function $g(b_t)$ is for theoretical illustration and is not imposed on the data.

[3]Equation (3) is derived from equations (1) and (2) employing the probabilities of the two underlying regimes (Survive or Collapse), denoted by q_t and $(1- q_t)$.

built-up of the bubble and expect the bubble to collapse, they cannot precisely estimate the time of the collapse. BK propose that abnormally high trading volume is a signal of changing market expectations about the future of a speculative bubble. We enrich the BK specification with a set of observable/macroeconomic variables, which we assume that investors monitor and which act as a signal of changing economic conditions and as a result changing market expectations that help them find the optimal exit-time from the market. Our set of signalling indicators comprises of variables typically employed in oil price determination.[4] Under this setting, we model the probability of survival as a function of both the bubble size and one of the indicators (z_t) as follows:

$$q_t = \Phi(\beta_{q_0} + \beta_{q_1} |b_t| + \beta_{q_2} z_t) \tag{8}$$

The model described in equations (4) - (8) is highly nonlinear and in order to linearise it we take the first order Taylor Series expansion of (4)-(6) around an arbitrary b_0 and z_0. The resulting linear regime switching model is given by the following set of equations:

$$\begin{gathered} BK - extended \\ R_{d,t+1} = \beta_{d0} + \varepsilon_{d,t+1}, \text{ where } \varepsilon_{d,t+1} \sim N(0, \sigma_d^2) \\ R_{s,t+1} = \beta_{s0} + \beta_{s1} b_t + \beta_{s2} z_t + \varepsilon_{s,t+1}, \text{ where } \varepsilon_{s,t+1} \sim N(0, \sigma_s^2) \\ R_{c,t+1} = \beta_{c0} + \beta_{c1} b_t + \varepsilon_{c,t+1}, \text{ where } \varepsilon_{c,t+1} \sim N(0, \sigma_c^2) \\ \Pr(State_{t+1} = D) = \eta_t = \Phi(\beta_{\eta_0} + \beta_{\eta_1} |b_t| + \beta_{\eta_2} s_t) \\ q_t = \Phi(\beta_{q_0} + \beta_{q_1} |b_t| + \beta_{q_2} z_t) \end{gathered} \tag{9}$$

The BK-extended model is estimated by maximising the associated likelihood function, given by the following formula:

$$\begin{aligned} &\prod_t [\{\eta_t\} \varphi \left(\frac{R_{t+1} - \beta_{d0}}{\sigma_d} \right) \sigma_d^{-1}] + \{1 - \eta_t\} \{q_t\} \varphi \left(\frac{R_{t+1} - \beta_{s0} - \beta_{s1} b_t - \beta_{s2} z_t}{\sigma_s} \right) \sigma_s^{-1} + \\ &\qquad + \{1 - \eta_t\} \{1 - q_t\} \varphi \left(\frac{R_{t+1} - \beta_{c0} - \beta_{c1} b_t}{\sigma_c} \right) \sigma_c^{-1}] \end{aligned} \tag{10}$$

where φ is the standard normal probability density function (pdf).

Obviously, when the signal indicator variable coincides with abnormal trading volume, the BK-extended specification coincides with the BK model. Furthermore, if we set the probability of being in the dormant regime equal to zero and exclude all explanatory

[4] Additional information about the variables-predictors we use in our analysis is provided in Subsection 2.3.

variables (other than b_t) from the specification, our model reduces to the original VNS model, which serves as the natural benchmark in our analysis and is outlined by the following set of equations.

$$\begin{gathered} VNS \\ R_{s,t+1} = \beta_{s0} + \beta_{s1} b_t + \varepsilon_{s,t+1}, \text{ where } \varepsilon_{s,t+1} \sim N(0, \sigma_s^2) \\ R_{c,t+1} = \beta_{c0} + \beta_{c1} b_t + \varepsilon_{c,t+1}, \text{ where } \varepsilon_{c,t+1} \sim N(0, \sigma_c^2) \\ \Pr(State_{t+1} = S) = q_t = \Phi(\beta_{q_0} + \beta_{q_1} |b_t|) \end{gathered} \tag{11}$$

To gauge the influence/significance of the candidate indicator variables, we also augment the VNS model by adding the signal variable z_t in both the mean equation for the survival regime and the probability of survive, q_t, as follows:

$$\begin{gathered} VNS - extended \\ R_{s,t+1} = \beta_{s0} + \beta_{s1} b_t + \beta_{s2} z_t + \varepsilon_{s,t+1}, \text{ where } \varepsilon_{s,t+1} \sim N(0, \sigma_s^2) \\ R_{c,t+1} = \beta_{c0} + \beta_{c1} b_t + \varepsilon_{c,t+1}, \text{ where } \varepsilon_{c,t+1} \sim N(0, \sigma_c^2) \\ \Pr(State_{t+1} = S) = q_t = \Phi(\beta_{q_0} + \beta_{q_1} |b_t| + \beta_{q_2} z_t) \end{gathered} \tag{12}$$

As previously, VNS and VNS-extended models are estimated by the associated likelihood functions.

2.2 Bubble calculation

The fundamental value of the oil price is defined as the sum of the current and expected convenience yield that accrues from holding inventories in the same way that the dividend yield is employed to estimate the fundamental value in the stock market. Following Pindyck (1993), SA and Lammerding *et al.* (2013), we use the futures prices in order to measure the convenience yield of actively traded storable commodities. More in detail, let $Y_{t,1}$ denote the current monthly convenience yield, r_t the risk-free interest rate, P_t the spot price of oil and F_{t+1} the futures price. Then, the convenience yield is calculated by the following non-arbitrage condition

$$Y_{t,1} = (1 + r_t) P_t - F_{t+1}$$

which simply states that in equilibrium the futures price must equal the spot price (adjusted for opportunity costs) and the benefits of holding the commodity. Once we obtain the convenience yield, we follow Campbell and Shiller (1987) in order to calculate the fundamental value and the bubble size by employing the convenience yield instead of the dividend yield. Specifically, the Campbell and Shiller model allows for predictable variation in expected convenience yields and is based on a simple present value model of stock prices. If the present value model of oil spot prices

$$P_t = E_t \sum_{i=0}^{\infty} (\frac{1}{1+r})^i Y_{t,i}$$

were true, then a linear function of current prices and convenience yields would be the optimal linear forecast of future convenience yields. This linear function, namely the spread (S_t), is defined as the difference between price and a multiple of current convenience yield and can be estimated through a VAR representation for the change in convenience yields and the spread itself. More in detail, the spread (S_t) and the implied bubble measure (b_t) are given by the following formulas

$$S_t = P_t^* - \frac{1+r}{r} D_t = \frac{1+r}{r} \sum_{i=0}^{\infty} (\frac{1}{1+r})^i E_t(\Delta D_{t+i})$$

$$b_t = 1 - \frac{S_t + ((1+r)/r) * D_{t-1}}{P_t}$$

We set the discount rate equal to

$$r = \frac{1}{\overline{P_t}/\overline{D_t} - 1}$$

to ensure that the spread has a zero mean over the whole sample.

2.3 Choice of indicator variables

In this section, we outline the rationale behind the choice of the signal/indicator variables employed in the extended BK and VNS models. As already mentioned, BK employ the abnormal futures trading volume as a signal of market expectations governing both the mean and the probability equation of the surviving regime. This variable can be thought of as a destabilising financial shock in the context of Lombardi and Van Robays (2011). Drawing from the structural vector autoregressive model of the global market for crude oil proposed by Kilian and Murphy (2013), we also employ variables linked to demand and supply shocks in the oil market. Since these models produce empirically plausible estimates of the impact of demand and supply shocks on the real price of oil, they may also have value for forecasting the real price of oil. The variables in this model include the percent change in global crude oil production, a measure of global real activity and the change in global above-ground crude oil inventories. Global crude oil production serves as the flow supply shock and global real activity, which is the dry cargo shipping rate index developed by Kilian (2009), serves as the flow demand shock and represents unexpected fluctuations in the global business cycle. Finally, the speculative demand shock, which is a shock to the demand for oil inventories arising from forward-looking behaviour of market participants is proxied by crude oil inventories. Widening the information set, we also employ macroeconomic and financial variables which act complementarily to measures of economic activity and financial conditions. First of all, following Juvenal and Petrella (2012), we consider an alternative measure of economic activity proxied by US industrial production. Admittedly, this variable is just a weak substitute for global economic activity, however the first factor the authors extract from a large dataset of 151 variables of the G7 countries loads primarily on industrial production. More importantly, this measure of economic activity is found to summarise aggregate business cycle conditions, while the proxy proposed by Kilian (2009) is more forward looking measure summarising aggregate demand and loading heavily on US personal consumption. The remaining variables we employ are mainly US variables and mainly contributing to the first factor with the exception of the US monetary base (M1) that is linked to the second factor. Anzuini *et al.* (2010) highlight that expansionary monetary policy may have fuelled oil price increases, but also report that it appears to exert its impact through expectations of higher inflation and growth rather than on the flow of global liquidity into oil futures markets. By no means is this list of variables exhaustive. For a detailed list of variables determining oil prices, the reader is referred to Alquist *et al.* (2011) and Juvenal and Petrella (2012).

3 Empirical findings

3.1 Data

Our dataset consists of monthly observations from January 1985 to December 2010. P_t and F_{t+1} are the prices of the nearest-month and the next-to-nearest month West Texas Intermediate (WTI) futures contract taken from the US Energy Information Administration, respectively. The risk free interest rate is the 3-month Treasury Bill rate and the US Consumer Price Index (CPI) is employed to deflate P_t and $Y_{t,1}$ (both taken from the Federal Reserve Economic Data, FRED). The futures trading volume data is based on nearest-month futures contracts of WTI (DataStream International)[5]. Following SA, abnormal trading volume is calculated as the percentage deviation of last months' volume from the 6-month moving average. Our measure of fluctuations in global real activity is the dry cargo shipping rate index developed in Kilian (2009). This real activity index is

[5]The data necessary for the calculation of the bubble along with the abnormal trading volume were kindly provided by Professor Shuping Shi.

a business cycle index, designed to capture shifts in the global use of industrial commodities, and stationary by construction.[6] Data on global crude oil production are available in the Monthly Energy Review of the Energy Information Administration (EIA). These data include lease condensates, but exclude natural gas plant liquids and are expressed in percent changes. Given the lack of data on crude oil inventories for other countries, we follow Hamilton (2009a) and Kilian and Murphy (2013) in using the data for total U.S. crude oil inventories provided by the EIA, scaled by the ratio of OECD petroleum stocks over U.S. petroleum stocks, also obtained from the EIA. We express the resulting proxy for global crude oil inventories in percent changes. Finally, the US macroeconomic/financial indicator variables are sourced from the FRED database. Specifically, we employ the long-term (DGS10) and short-term interest rate (GS3M), the growth rate in industrial production (INDPRO), the CPI (CPIAUCSL) and PPI (PPIACO) inflation rate, the growth in M1 money stock (M1) and the unemployment rate (UNRATE).

Figure 1 presents the price of oil together with the calculated deviations from fundamental values. We observe significant negative deviations in the late '80s and late '90s. On the other hand, there seems to be a noteworthy positive bubble in the last two years of our sample.

[FIGURE 1 AROUND HERE]

3.2 Forecasting - Statistical evaluation

We focus on one-period ahead forecasts and organise the forecast exercise in real-time terms, i.e. we obtain a forecast for period $t + 1$ using all available information up to period t. All models and the deviations from fundamentals are re-estimated recursively and the first estimation sample ends in December 2002, leaving the last 8 years of our sample (about one third of the available observations) for the out-of-sample evaluation period.

Table 1 reports the ratio of the Mean Squared Forecast Error (MSFE) of each one of the RS models over the MSFE of the RW model.[7] A ratio below unity indicates the superiority of our RS model over the RW model. In all cases, the RS models outperform the benchmark and the MSFE ratios are usually well below unity. In the case of the simple VNS model that does not contain any explanatory variables (other than the deviations from fundamentals), the MSFE ratio equals 0.929. When we enrich the specification of the RS model with one of the twelve predictors considered in our analysis (VNS-

[6]The index is available from Lutz Kilian's website: http://www-personal.umich.edu/~lkilian/paperlinks.html

[7]We consider both the RW model with and without a drift. The former produces more accurate forecast than the latter. We choose to compare our RS models with the optimal RW model (i.e. the one that includes a drift).

extended model), the MSFE ratios range from 0.89 (CPI inflation rate) to 0.98 (world oil production) depending on the predictor. Substantial improvements in the forecasting performance of the RS models are observed when we allow for a third regime in our specification (BK-extended model). In this case, the MSFE ratios are usually around 0.9 with a minimum value of 0.856 when we use the industrial production as a predictor.

Among the twelve variables considered in our analysis, the growth rate of industrial production and the inflation rate seem to be the optimal predictors (in the context of our RS models) for the price of oil. Comparing the relative forecasting accuracy of the VNS-extended and the BK-extended models, we observe that, in general, the three-state RS models produce lower MSFEs than the two-state ones.

[TABLE 1 AROUND HERE]

Turning to the statistical significance of the predictive power of our models and given that we are interested in comparing the forecasting performance of nested models, we apply the methodology developed by Clark and West (2006, 2007). Specifically, let $R_{RW,t}$ and $R_{RS,t}$ denote the forecasts for R_t obtained from the RW and RS model respectively. Given a sequence of P forecasts, we first calculate:

$$f_t = (R_t - R_{RW,t})^2 - (R_t - R_{RS,t})^2 + (R_{RW,t} - R_{RS,t})^2, t = 1, 2, ..., P$$

The test statistic of Clark and West is given by the standard $t-statistic$ of the regression of f_t on a constant. Clark and West (2006, 2007) recommend using 1.282 and 1.645 for a 0.10 and 0.05 test, respectively. We should note that this is a one-sided test. Clark and West show that under the null hypothesis of equal MSFE, the unrestricted model (the RS model in our case) should generate larger MSFE than the restricted one (RW in our case). The intuition behind this argument is that since under the null hypothesis the additional parameters of the unrestricted model do not help predictions, in finite samples this model loses efficiency due to the estimation of these parameters that introduces noise into the forecasts. This inflates the MSFE of the model. Therefore, even if the restricted model generates smaller MSFE than the unrestricted one, we should not consider this as prima facie evidence of superiority of the former over the latter.

The findings are presented in Table 1 where the asterisks next to the calculated MSFE ratio denote rejection of the null hypothesis at the relevant confidence level. The results clearly suggest that the superiority of the RS model relative to the RW model is always statistically significant. We go one step further and examine whether the inclusion of a predictor, z_t, in our RS models improves the accuracy of forecasts by comparing the predictive power of VNS-extended and BK-extended relative to that of the simple VNS model that contains no predictors. Entries in bold in Table 1 indicate cases where the extended RS model generates forecasts that are statistically more accurate compared to those from the simple VNS model. Both the VNS-extended model and the BK-extended model outperform the simple VNS model in five out of twelve cases. Among the predictors under scrutiny, the CPI and PPI inflation rates achieve the best forecasting performance of the RS models considered in this study.

3.3 Economic evaluation

We also perform an economic evaluation of the forecasts of our RS models based on (i) the utility-based approach initiated by West *et al.* (1993) and (ii) the manipulation-proof performance measure proposed by Goetzmann *et al.* (2007). We briefly describe the two approaches. Consider a US investor who dynamically rebalances her portfolio every month. Her portfolio choice problem is how to allocate wealth between a risk-free asset yielding interest rate i_t,[8] and the risky future contract on the price of oil. In a mean-variance framework and given a specific Relative Risk Aversion (RRA) coefficient, γ, that controls the investor's appetite for risk (Campbell and Viceira, 2002), she chooses the optimal weight on the risky asset in a standard maximisation problem resulting in a portfolio return over the out-of-sample period equal to, say, $r_{p,t+1}$. Then, over the forecast evaluation period the investor with initial wealth of $W_o = 1\$$ realises an average utility of

$$\overline{U} = \frac{1}{P}\sum_{t=0}^{P-1}(r_{p,t+1} - \frac{\gamma}{2}(r_{t+1} - \overline{r}_{t+1})^2) \tag{13}$$

A risk-averse investor will be willing to pay a performance fee, denoted by Φ, for switching from the portfolio constructed based on RW forecasts to a portfolio based on our proposed RS forecasts if the latter are superior to the former. In our set-up the performance fee is calculated as the difference between the realised utilities as follows:

$$\begin{aligned} \frac{1}{P}\sum_{t=0}^{P-1}(r^{RS}_{p,t+1} - \Phi - \frac{\gamma}{2}(r_{t+1} - \overline{r}_{t+1})^2) &= \frac{1}{P}\sum_{t=0}^{P-1}(r^{RW}_{p,t+1} - \frac{\gamma}{2}(r_{t+1} - \overline{r}_{t+1})^2) \Longrightarrow \\ \Phi &= \overline{U}^{RS} - \overline{U}^{RW} \end{aligned} \tag{14}$$

Positive values of Φ suggest superior predictive ability of the RS model against the RW benchmark.

On the other hand, the manipulation-proof performance measure, $M(R_p)$, can be interpreted as a portfolio's premium return after adjusting for risk and it remedies potential caveats associated with the popular Sharpe ratio such as the effect of non-normality, the

[8] We employ the 1-month US Certificate of Deposit (CD1M, FRED).

underestimation of the performance of dynamic strategies and the choice of utility function. The difference, Θ, between the $M(R_p)s$ of competing models is employed to assess the most valuable model. This measure is defined as

$$\begin{aligned} M(R_p) &= \frac{1}{1-\gamma} \ln \left\{ \frac{1}{P} \sum_{t=0}^{P-1} \left(\frac{1 + r_{p,t+1}}{i_t} \right)^{1-\gamma} \right\} \\ \Theta &= M(R_p)^{RS} - M(R_p)^{RW} \end{aligned} \tag{15}$$

Table 2 reports both economic evaluation measures in monthly percentage points. We do not allow short-selling and we consider two different values for the RRA coefficient ($\gamma = 3$ and $\gamma = 7$). In general, both measures produce qualitatively similar results. We always observe significant gains in economic terms for the investor who is willing to switch from the benchmark RW model to one of the RS models. For example, when $\gamma = 3$, the utility gain Φ equals 0.42% for the simple VNS model, while the average utility gains for the VNS-extended and the BK-extended models are 0.71% and 0.69%, respectively. We observe the highest utility gain in the case of the three-sate model that includes the industrial production as a predictor. In this case, Φ equals 1.57% which corresponds to a 18.84% gain in annual terms. Contrary to our findings for the statistical evaluation of our forecasts, the VNS-extended model seems to generate slightly higher utility gains relative to the BK-extended model. It is also interesting to note that models with low MSFE ratios do not necessarily generate higher gains in economic terms. For example, in the case of the three-state model that includes the world oil production, the MSFE ratio relative to the RW model is as low as 0.896, while both economic evaluation measures are pretty low ($\gamma = 3$). In general, as the RRA coefficient increases, utility gains of the RS models relative to the RW model become more pronounced.

[TABLE 2 AROUND HERE]

Finally, we repeat the economic evaluation procedure replacing RW with VNS. In other words, we compare VNS-extended and BK-extended with the VNS model. In this way, we investigate whether the inclusion of a predictor in the RS specification produces any benefits in economic terms. These findings are reported in Table 3. In the majority of cases, we observe positive values of Φ and Θ, reaching 1.15% and 1.20%, respectively, in the case of the BK-extended model that includes the growth rate of the industrial production as a predictor ($\gamma = 3$). In general, the benefits from VNS-extended are higher compared to those from BK-extended.

[TABLE 3 AROUND HERE]

4 Conclusions

We develop two- and three-state Regime-Switching (RS) models and test their forecasting ability for oil prices. Taking advantage of the deviations we periodically observe between the market price of oil and its fundamental value, our models relate the expected gross return in the oil price to deviations from fundamentals and an additional explanatory variable. Specifically, we compare the predictive power (in both statistical and economic evaluation terms) of twelve alternative macroeconomic/indicator variables assuming a forecast horizon of one month. Our findings indicate substantial benefits, in terms of forecasting accuracy, when RS models are employed relative to the Random Walk (RW) benchmark, especially in economic evaluation terms. Moreover, the RS models enriched with one of the predictors proposed in this study often outperform simple RS models that contain no predictors (other than deviations from fundamentals).

Our findings reveal that speculative behaviour models can be used to generate reliable out-of-sample forecasts for the price of oil. The analysis opens routes for future research to other commodities, such as gold and wheat. It would also be very interesting to investigate whether combination of forecasts from various models can lead to additional accuracy gains.

References

[1] Alquist, R., Kilian, L., 2010. What do we learn from the price of crude oil futures?, *Journal of Applied Econometrics*, 25, 539-573.

[2] Alquist, R., Kilian, L., Vigfusson, R.J., 2011. Forecasting the price of oil, forthcoming in: G. Elliott and A. Timmermann (eds.), *Handbook of Economic Forecasting*, 2, Amsterdam: North-Holland.

[3] Anzuini, A., Lombardi, M.J., Pagano, P., 2010. The impact of monetary policy shocks on commodity prices, *ECB Working Paper* 1232.

[4] Baumeister, C., Kilian, L., 2012. Real-time forecasts of the real price of oil, *Journal of Business and Economic Statistics*, 30(2), 326-336.

[5] Baumeister, C., Kilian, L., 2013. What central bankers need to know about forecasting oil prices, *Bank of Canada*, Working Paper 2013-15.

[6] Branch, W.A., Evans, G.W., 2011. Learning about risk and return: a simple model of bubbles and crashes, *American Economic Journal: Macroeconomics* 3, 159-191.

[7] Brooks, C., Katsaris, A., 2005. A three-regime model of speculative behaviour: Modelling the evolution of the S&P 500 composite index, *The Economic Journal*, 115, 767–797.

[8] Campbell, J.Y., Shiller, R.J., 1987. Cointegration and tests of present value models, *The Journal of Political Economy* 95, 1062–1088.

[9] Campbell, J.Y., Viceira, L., 2002. Strategic asset allocation, Oxford: *Oxford University Press.*

[10] Clark, T.E., West, K.D., 2006. Using out-of-sample mean squared prediction errors to test the martingale difference hypothesis, *Journal of Econometrics*, 135, 155-186.

[11] Clark, T.E., West, K.D., 2007. Approximately normal tests for equal predictive accuracy in nested models, *Journal of Econometrics*, 138 (1), 291–311.

[12] CFTC, 2011. CFTC commitments of traders, *US commodity futures trading commission*, Washington, DC.

[13] Einloth, J.T., 2009. Speculation and recent volatility in the price of oil, *Social Science Research Network* (SSRN) Working Paper.

[14] Evans, G.W., 1991. Pitfalls in testing for explosive bubbles in asset prices, *American Economic Review,* 81, 922-930.

[15] Fattouh, B., Kilian, L., Mahadeva, L., 2013. The role of speculation in oil markets: What have we learned so far?, *Energy Journal*, 34 (3), forthcoming.

[16] Gilbert, C.L., 2010. Speculative influences on commodity futures prices 2006–2008, *United Nations Conference on Trade and Development* (UNCTAD) Discussion Paper No. 197.

[17]Goetzmann, W., Ingersoll, J., Spiegel, M., Welch, I., 2007. Portfolio performance manipulation and manipulation-proof performance measures, *Review of Financial Studies*, 20, 1503-1546.

[18]Gorton, G., Rouwenhorst, K.G, 2006. Facts and fantasies about commodity futures, *Financial Analysts Journal*, 62(02), 47–68.

[19]Hamilton, J.D., 2009a. Causes and consequences of the oil shock of 2007–2008, *Brookings Papers on Economic Activity*, 215-259.

[20]Hamilton, J.D., 2009b. Understanding crude oil prices, *Energy Journal*, 30 (2), 179–206.

[21]Homm, U., Breitung, J., 2012. Testing for speculative bubbles in stock markets: A comparison of alternative methods. *Journal of Financial Economics*, 10 (1), 198–231.

[22]Irwin, S.H., Sanders, D.R., 2011. Index funds, financialization and commodity futures markets, *Applied Economic Perspectives and Policy,* 33, 1-31.

[23]Juvenal, L., Petrella, I., 2012. Speculation in the oil markets, Working Paper 2011-027E (revised June 2012), *Federal Reserve Bank of St. Louis*, Working Paper Series.

[24]Kilian, L., 2009. Not all oil price shocks are alike: disentangling demand and supply shocks in the crude oil market, *American Economic Review,* 99 (3), 1053–1069.

[25]Kilian, L., Hicks, B., 2009. Did unexpectedly strong economic growth cause the oil price shock of 2003–2008?, *Journal of Forecasting*, forthcoming.

[26]Kilian, L., Lee, T.K., 2013. Quantifying the speculative component in the real price of oil: The role of global oil inventories, *Journal of International Money and Finance*, forthcoming.

[27]Kilian, L., Murphy, D., 2013. The role of inventories and speculative trading in the global market for crude oil, *Journal of Applied Econometrics*, forthcoming.

[28]Knetsch, T.A., 2007. Forecasting the price of oil via convenience yield predictions, *Journal of Forecasting*, 26, 527-549.

[29]Lammerding, M., Stephan, P, Trede, M., Wilfling, B., 2013. Speculative bubbles in recent oil price dynamics: Evidence from a Bayesian Markov-switching state-space approach, *Energy Economics*, 36, 491-502.

[30]Lombardi, M.J., Van Robays, I., 2011. Do financial investors destabilize the oil price?, Working Paper Series No. 1346. *European Central Bank,* Frankfurt/Main.

[31]Mayer, J, 2010. The financialization of commodity markets and commodity price volatility, In Dullien, S., Kotte, D.J., Marquez, A. Priewe, J. (Eds.) *The Financial and Economics Crisis of 2008-2009 and Developing Countries*. UNCTAD, New York and Geneva.

[32]Pindyck, R.S., 1993. The present value model of rational commodity pricing, *The Economic Journal*, 103, 511-530.

[33]Phillips, P.C.B., Yu, J., 2011. Dating the timeline of financial bubbles during the subprime crisis, *Quantitative Economics,* 2 (3), 455–491.

[34]Phillips, P.C.B., Wu, Y., Yu, J., 2011. Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values?, *International Economic Review*, 52, 201-226.

[35]Shi, S., Arora, V., 2012. An application of models of speculative behaviour to oil prices, *Economics Letters* 115, 469–472.

[36]Tang, K., Xiong, W., 2011, Index investment and financialization of commodities, *Financial Analysts Journal* 68 (6), 54-74

[37]van Norden S., Schaller, H., 1993. The predictability of stock market regime: Evidence from the Toronto stock exchange, *Review of Economics and Statistics* 75, 3, 505–510.

[38]Went, P., Jirasakuldech, B., Emekter, R., 2012. Rational speculative bubbles and commodities markets: Application of duration dependence test, *Applied Financial Economics,* 22 (7), 581–596.

[39]West, K., Edison, H., Cho, D., 1993. A utility-based comparison of some models of exchange rate volatility, *Journal of International Economics*, 35, 23-46.

[40]Wu, T., McCallum, A., 2005. Do Oil Futures Prices Help Predict Future Oil Prices?, *Federal Reserve Bank of San Francisco Economic Letter*, 2005-38.

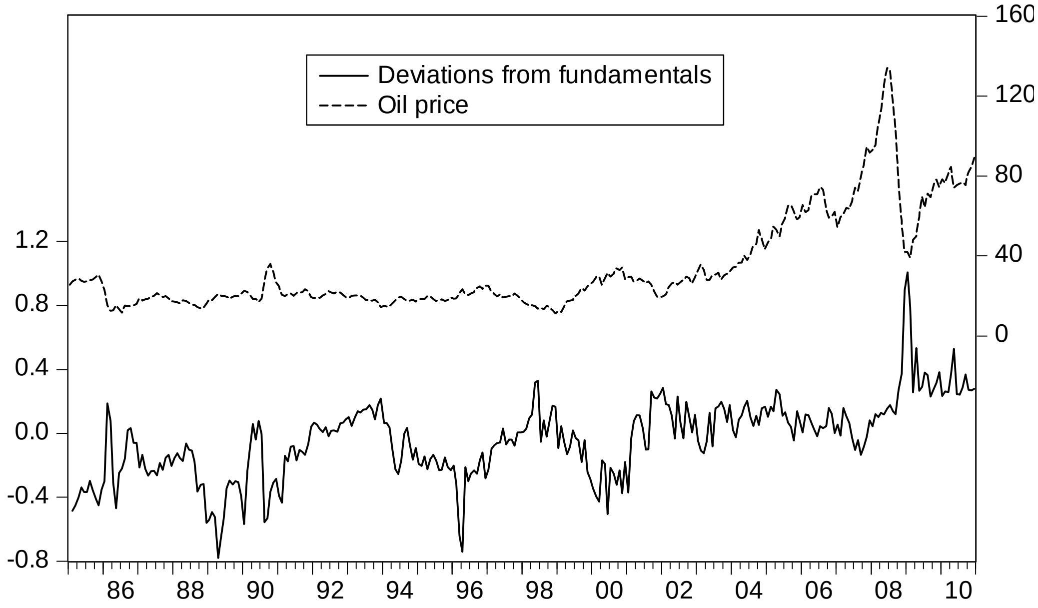

Figure 1. Deviations from fundamentals (left axis) and the price of oil (right axis)

Table 1. Mean Squared Forecast Errors (MSFEs) Ratios and Statistical Evaluation

VNS = 0.929**

	VNS-extended	**BK-extended**
Global real activity index	**0.968****	0.912**
Global real activity (% change)	**0.909****	0.930**
Inventories (% change)	0.960**	0.990*
World oil production	0.980**	**0.896****
Volume	**0.914****	0.950*
Long-term interest rate	0.961**	**0.908****
Industrial production	0.920**	**0.856****
CPI inflation rate	**0.890****	**0.918****
M1	0.976**	0.907**
PPI inflation rate	**0.951****	**0.877****
Short-term interest rate	0.956**	0.939**
Unemployment rate	0.930**	0.908**

Notes: The table reports the ratio of the Mean Squared Forecast Error (MSFE) of the Regime-Switching (RS) model over the MSFE of the Random Walk (RW) model. (**) and (*) denote the superiority in statistical terms of the RS model relative to the RW model at a 5% and 10% confidence level respectively. Entries in bold indicate the superiority in statistical terms of the corresponding model relative to the VNS model at a 10% confidence level.

Table 2. Economic Evaluation (relative to the Random Walk model)

	$\gamma = 3$		$\gamma = 7$	
	Φ	Θ	Φ	Θ
VNS	0.42	0.58	0.92	1.68
VNS-extended				
Global real activity index	0.80	0.97	1.09	1.80
Global real activity (% change)	0.56	0.73	0.99	1.73
Inventories (% change)	0.58	0.74	0.92	1.64
World oil production	0.75	0.92	1.04	1.79
Volume	0.70	0.87	1.07	1.79
Long-term interest rate	0.34	0.53	0.79	1.54
Industrial production	0.67	0.84	0.99	1.71
CPI inflation rate	1.22	1.45	1.23	2.03
M1	0.80	0.97	1.06	1.78
PPI inflation rate	1.01	1.18	1.09	1.83
Short-term interest rate	0.63	0.81	0.96	1.69
Unemployment rate	0.41	0.59	0.87	1.63
BK-extended				
Global real activity index	0.70	0.82	0.68	1.04
Global real activity (% change)	0.41	0.52	0.92	1.54
Inventories (% change)	0.70	0.87	0.59	0.98
World oil production	0.33	0.34	0.89	1.43
Volume	0.23	0.36	0.80	1.49
Long-term interest rate	0.89	1.07	1.42	2.23
Industrial production	1.57	1.78	1.90	2.68
CPI inflation rate	0.84	0.95	1.13	1.78
M1	0.56	0.65	0.83	1.28
PPI inflation rate	0.92	1.09	1.38	2.14
Short-term interest rate	0.77	0.89	0.66	1.18
Unemployment rate	0.32	0.33	0.74	1.35

Notes: γ stands for the Relative Risk Aversion coefficient. Entries correspond to the value of Φ and Θ when we compare the Regime-Switching (RS) model relative to the Random Walk (RW) model. The performance fee, Φ, is the fraction of the wealth which when subtracted from the RS proposed portfolio returns equates the average utilities of the competing model (i.e. the RS and the RW models). If our proposed RS model does not contain any economic value, the performance fee is negative ($\Phi \preceq 0$), while positive values of the performance fee suggest superior predictive ability against the RW benchmark. Θ is the difference between the manipulation-proof performance measure of competing models (RS and RW). Both Φ and Θ and are reported in percentage points on a monthly basis.

Table 3. Economic Evaluation (relative to the VNS model)

	$\gamma = \mathbf{3}$		$\gamma = \mathbf{7}$	
	Φ	Θ	Φ	Θ
VNS-extended				
Global real activity index	0.37	0.39	0.17	0.12
Global real activity (% change)	0.14	0.15	0.07	0.05
Inventories (% change)	0.16	0.16	-0.01	-0.03
World oil production	0.33	0.33	0.12	0.11
Volume	0.28	0.29	0.14	0.11
Long-term interest rate	-0.08	-0.05	-0.14	-0.14
Industrial production	0.25	0.25	0.06	0.03
CPI inflation rate	0.80	0.86	0.31	0.35
M1	0.37	0.39	0.13	0.11
PPI inflation rate	0.59	0.60	0.17	0.16
Short-term interest rate	0.21	0.22	0.04	0.02
Unemployment rate	-0.01	0.00	-0.05	-0.05
BK-extended				
Global real activity index	0.28	0.24	-0.25	-0.64
Global real activity (% change)	-0.01	-0.07	0.00	-0.14
Inventories (% change)	0.28	0.29	-0.33	-0.69
World oil production	-0.09	-0.25	-0.03	-0.25
Volume	-0.19	-0.22	-0.12	-0.18
Long-term interest rate	0.47	0.49	0.50	0.55
Industrial production	1.15	1.20	0.98	1.01
CPI inflation rate	0.42	0.36	0.21	0.10
M1	0.14	0.06	-0.10	-0.39
PPI inflation rate	0.50	0.50	0.45	0.47
Short-term interest rate	0.35	0.31	-0.26	-0.50
Unemployment rate	-0.11	-0.25	-0.18	-0.33

Notes: γ stands for the Relative Risk Aversion coefficient. Entries correspond to the value of Φ and Θ when we compare either the VNS-extended or the BK-extended model relative to the simple VNS model. The performance fee, Φ, is the fraction of the wealth which when subtracted from the extended RS proposed portfolio returns equates the average utilities of the competing model (i.e. the extended RS and the simple VNS models). If the extended RS model does not contain any economic value, the performance fee is negative ($\Phi \preceq 0$), while positive values of the performance fee suggest superior predictive ability against the VNS benchmark. Θ is the difference between the manipulation-proof performance measure of competing models (extended RS and VNS). Both Φ and Θ and are reported in percentage points on a monthly basis.