Quantifying time-varying uncertainty and risk for the real price of oil

Herman K. van Dijk[1]

[1]Erasmus University, Tinbergen Institute and Norges Bank

Vienna University of Economics and Business
Research Seminar of the Institute for Statistics and Mathematics
October 19, 2021

Disclaimer: The views expressed herein are solely those of the presenter and do not necessarily reflect the views of Norges Bank. Background is paper: **Quantifying time-varying uncertainty and risk for the real price of oil**. Joint work with **Knut Are Aasveit and Jamie Cross** from Norges Bank and BI Norwegian Business School

General motivation

- Oil price fluctuations have **adverse** macroeconomic implications
  - Hamilton (1983, 2009); Kilian (2009); Ravazzolo and Rothman (2013); Baumeister and Hamilton (2019)

General motivation

- Oil price fluctuations have **adverse** macroeconomic implications
  - Hamilton (1983, 2009); Kilian (2009); Ravazzolo and Rothman (2013); Baumeister and Hamilton (2019)
- Oil price is a **key** variable in generating macroeconomic projections and in assessing macroeconomic risks
  - **Central banks, private sector forecasters and international organizations**

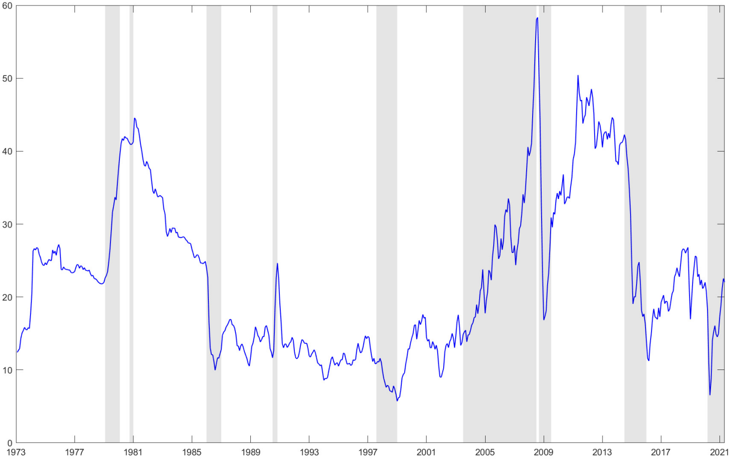
General motivation

- Oil price fluctuations have **adverse** macroeconomic implications
  - Hamilton (1983, 2009); Kilian (2009); Ravazzolo and Rothman (2013); Baumeister and Hamilton (2019)
- Oil price is a **key** variable in generating macroeconomic projections and in assessing macroeconomic risks
  - **Central banks, private sector forecasters and international organizations**
- But also crucial for how some sectors operate their business
  - **Airlines, utilities and automobile manufacturers**

General motivation

- Oil price fluctuations have **adverse** macroeconomic implications
  - Hamilton (1983, 2009); Kilian (2009); Ravazzolo and Rothman (2013); Baumeister and Hamilton (2019)
- Oil price is a **key** variable in generating macroeconomic projections and in assessing macroeconomic risks
  - **Central banks, private sector forecasters and international organizations**
- But also crucial for how some sectors operate their business
  - **Airlines, utilities and automobile manufacturers**
- ...But the **price of oil is not easy to forecast**

Specific Motivation 1: Given the changing data pattern in the real price of oil which is subjected to shocks and volatility, how to forecast it ?



More on Motivation 1

- Hamilton (2009) documents that the **statistical regularities of changes in the real price of oil** have historically tended to be

More on Motivation 1

- Hamilton (2009) documents that the **statistical regularities of changes in the real price of oil** have historically tended to be
  - **permanent**
  - **difficult to predict**
  - governed by **very different regimes at different points in time**

More on Motivation 1

- Hamilton (2009) documents that the **statistical regularities of changes in the real price of oil** have historically tended to be
  - **permanent**
  - **difficult to predict**
  - governed by **very different regimes at different points in time**
- He further argues that the price of oil seems to follow a **random walk** without drift

More on Motivation 1

- Hamilton (2009) documents that the **statistical regularities of changes in the real price of oil** have historically tended to be
  - **permanent**
  - **difficult to predict**
  - governed by **very different regimes at different points in time**

- He further argues that the price of oil seems to follow a **random walk** without drift

- It is widely accepted to either use the **current spot price or the price of oil futures contracts** as the forecast of the price of oil.

Oil Price forecasting literature: point forecasting

- Recently academic and professional researchers have explored numerous alternative models and methods in order to forecast the most likely future realisation of the oil price

Oil Price forecasting literature: point forecasting

- Recently academic and professional researchers have explored numerous alternative models and methods in order to forecast the most likely future realisation of the oil price
  - Alquist and Kilian (2010); Alquist et al. (2013); Baumeister and Kilian (2012, 2015); Manescu and Robays (2016); Bernard et al. (2018); Pak (2018); Garratt et al. (2019); Baumeister et al. (2020)

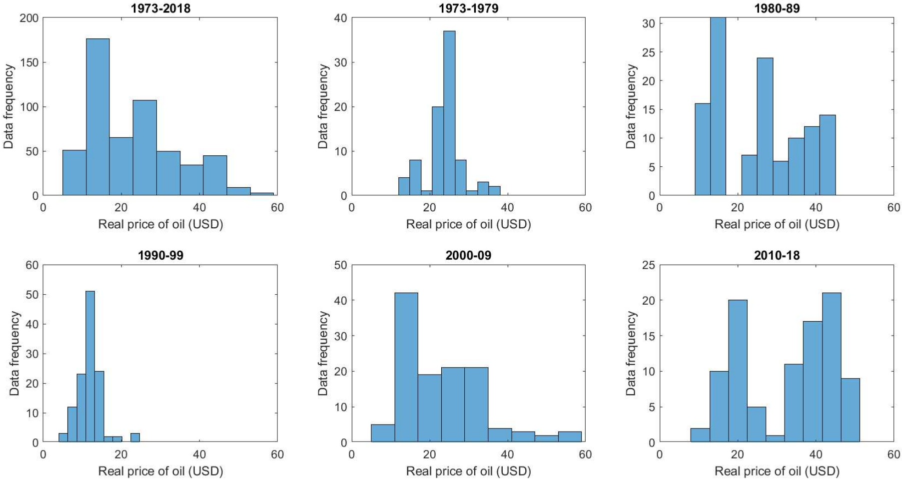
Oil Price forecasting literature: point forecasting

- Recently academic and professional researchers have explored numerous alternative models and methods in order to forecast the most likely future realisation of the oil price
  - Alquist and Kilian (2010); Alquist et al. (2013); Baumeister and Kilian (2012, 2015); Manescu and Robays (2016); Bernard et al. (2018); Pak (2018); Garratt et al. (2019); Baumeister et al. (2020)
- These papers focus on evaluating **point forests** and find that

Oil Price forecasting literature: point forecasting

- Recently academic and professional researchers have explored numerous alternative models and methods in order to forecast the most likely future realisation of the oil price
  - Alquist and Kilian (2010); Alquist et al. (2013); Baumeister and Kilian (2012, 2015); Manescu and Robays (2016); Bernard et al. (2018); Pak (2018); Garratt et al. (2019); Baumeister et al. (2020)
- These papers focus on evaluating **point forests** and find that
  - It's **hard to beat a random walk** in out-of-sample oil price forecasting exercises
  - But careful attention to the **economic fundamentals** that are driving energy markets can lead to practical improvements in forecasts

Specific Motivation 2: How to model and forecast the changing distribution of real price of oil ?



Promising approach: Combine information from different sources

Basic practice of combining information in macroeconomic and financial forecasting is to make use of a **weighted combination** of **forecasts** from many sources, say **experts, models and/or large micro-data sets.** Let y_t be the variable of interest, and let $\tilde{y}_{1t}, \ldots, \tilde{y}_{n,t}$ be forecasted values from $i = 1, ..., n$ models, with weights $w_{1t}, \ldots, w_{n,t}$ where n maybe small (from a committee of experts) or large (from a large micro-data set). Then, basic practice is to make use of:

$$y_t = \sum_{i=1}^{n} w_{it} \tilde{y}_{it} \tag{1}$$

where $\tilde{y}_{it}$ should be a good approximation to y_t.

- **Problem with Practice** Many agencies handle this averaging and updating informally.

Our approach in combining information from different sources

- **Challenge** Give this practice a **Bayesian probabilistic foundation** in order to evaluate **practical** issues as follows: Make use of **Forecast Density Combinations (FDC)** features which, given information on data en model specification allows the evaluation of **Conditional Probabilities of (extreme) events: Recession probability; Turning point probability; Probabilistic warnings about defaults and crises in macroeconomics and finance; Value-at-Risk etc etc.**

Our approach in combining information from different sources

- **Challenge** Give this practice a **Bayesian probabilistic foundation** in order to evaluate **practical** issues as follows: Make use of **Forecast Density Combinations (FDC)** features which, given information on data en model specification allows the evaluation of **Conditional Probabilities of (extreme) events: Recession probability; Turning point probability; Probabilistic warnings about defaults and crises in macroeconomics and finance; Value-at-Risk etc etc.**
- **Fast growth in Big Data** gives opportunity of more accurate forecast measures. Analogy with weather forecasting using many satellite pictures. But in economics issue like multimodality, skewness, diverging ratio's of, for instance, government expenditures and GDP are not trivial and time-varying.

Our approach in combining information from different sources

- **Challenge** Give this practice a **Bayesian probabilistic foundation** in order to evaluate **practical** issues as follows: Make use of **Forecast Density Combinations (FDC)** features which, given information on data en model specification allows the evaluation of **Conditional Probabilities of (extreme) events: Recession probability; Turning point probability; Probabilistic warnings about defaults and crises in macroeconomics and finance; Value-at-Risk etc etc.**
- **Fast growth in Big Data** gives opportunity of more accurate forecast measures. Analogy with weather forecasting using many satellite pictures. But in economics issue like multimodality, skewness, diverging ratio's of, for instance, government expenditures and GDP are not trivial and time-varying.
- **New Tool:Parallel Computing: New Hardware and Software** give openings to solve complex problems. **Machine learning with several hidden layers using neural networks have a direct connection with filtering methods in nonlinear time series models**.

Relation to literature on combining probabilistic forecasts

- Combining forecast densities using **weighted linear combinations** of prediction models, evaluated using various scoring rules
  - Hall and Mitchell (2007); Amisano and Giacomini (2007); Jore et al. (2010); Hoogerheide et al. (2010); Kascha and Ravazzolo (2010); Geweke and Amisano (2011, 2012); Gneiting and Ranjan (2013); Aastveit et al. (2014)

Relation to literature on combining probabilistic forecasts

- Combining forecast densities using **weighted linear combinations** of prediction models, evaluated using various scoring rules
  - Hall and Mitchell (2007); Amisano and Giacomini (2007); Jore et al. (2010); Hoogerheide et al. (2010); Kascha and Ravazzolo (2010); Geweke and Amisano (2011, 2012); Gneiting and Ranjan (2013); Aastveit et al. (2014)

- **Complex combination approaches that allows for time-varying weights with possibly both learning and model set incompleteness**
  - Koop and Korobilis (2012); Billio et al. (2013); Casarin et al. (2015); Pettenuzzo and Ravazzolo (2016); Del Negro et al. (2016); Aastveit et al. (2018); McAlinn and West (2019); McAlinn et al. (2020); Takanashi and McAlinn (2020); Casarin et al. (2020)

Relation to literature on combining probabilistic forecasts

- Combining forecast densities using **weighted linear combinations** of prediction models, evaluated using various scoring rules
  - Hall and Mitchell (2007); Amisano and Giacomini (2007); Jore et al. (2010); Hoogerheide et al. (2010); Kascha and Ravazzolo (2010); Geweke and Amisano (2011, 2012); Gneiting and Ranjan (2013); Aastveit et al. (2014)

- **Complex combination approaches that allows for time-varying weights with possibly both learning and model set incompleteness**
  - Koop and Korobilis (2012); Billio et al. (2013); Casarin et al. (2015); Pettenuzzo and Ravazzolo (2016); Del Negro et al. (2016); Aastveit et al. (2018); McAlinn and West (2019); McAlinn et al. (2020); Takanashi and McAlinn (2020); Casarin et al. (2020)

- **No studies on how to quantify forecast uncertainty associated with the dynamic behaviour of the real price of crude oil.**

Plan of talk

- **Theoretical contributions: Structure of Forecast Density Combination (FDC): Probability Model, Equation System and Algorithm**
- Choice of model set for our empirical application
- **Empirical Contributions: Working of FDC in measuring time-varying uncertainty and risk in the real price of oil**
- Conclusion

Theoretical Contributions: Summary

- **Flexible Bayesian Forecast Density Combination** allows for cross-section and time dependent **Bayesian weight learning** and **Diagnostic learning about model incompleteness**. This **self-learning** is closely related to **Machine-Learning**.

Theoretical Contributions: Summary

- **Flexible Bayesian Forecast Density Combination** allows for cross-section and time dependent **Bayesian weight learning** and **Diagnostic learning about model incompleteness**. This **self-learning** is closely related to **Machine-Learning**.
- **Model Representation and Efficient Computation** Model is a **Generalised Linear State Space Model which allows the use of numerically efficient standard Markov Chain Monte Carlo simulation methods**. **Filtering** methods from State Space models are directly connected to the **integration of the hidden layers** in machine learning.

Structure of our FDC: Probability Model

- Let $\tilde{\boldsymbol{y}}_{\boldsymbol{t}}' = (\tilde{y}_{1t}, \ldots, \tilde{y}_{nt})$ be the forecasted values from $i = 1, ..., n$ models for the variable of interest y_t. In a simulation context $\tilde{y}_{it}$ **is a draw from the forecast distribution of model** M_i **with density** $p(\tilde{y}_{it}|I_{it-1}, M_i)$ and data set I_{it-1}.

Structure of our FDC: Probability Model

- Let $\tilde{\boldsymbol{y}}_{\boldsymbol{t}}{}' = (\tilde{y}_{1t}, \ldots, \tilde{y}_{nt})$ be the forecasted values from $i = 1, ..., n$ models for the variable of interest y_t. In a simulation context $\tilde{y}_{it}$ **is a draw from the forecast distribution of model** M_i **with density** $p(\tilde{y_{it}}|I_{it-1}, M_i)$ and data set I_{it-1}.
- Let $\boldsymbol{v}_{\boldsymbol{t}}{}' = (v_{0t}, v_{1t}, \ldots, v_{nt})$ **be latent continuous random variable parameters which are used to weight the different forecasts and combine these forecasts**

Structure of our FDC: Probability Model

- Let $\tilde{\boldsymbol{y}}_{\boldsymbol{t}}{}' = (\tilde{y}_{1t}, \ldots, \tilde{y}_{nt})$ be the forecasted values from $i = 1, ..., n$ models for the variable of interest y_t. In a simulation context $\tilde{y}_{it}$ **is a draw from the forecast distribution of model** M_i **with density** $p(\tilde{y_{it}}|I_{it-1}, M_i)$ and data set I_{it-1}.
- Let $\boldsymbol{v_t}' = (v_{0t}, v_{1t}, \ldots, v_{nt})$ **be latent continuous random variable parameters which are used to weight the different forecasts and combine these forecasts**
- The decomposition of the joint density of $(y_t, \boldsymbol{v}_t, \tilde{\boldsymbol{y}}_t)$ for the case of continuous random variables is:

$$p(y_t|I_{t-1}, M) = \int\int p(y_t|\boldsymbol{v}_t, \tilde{\boldsymbol{y}}_{\boldsymbol{t}})p(\boldsymbol{v}_t|\tilde{\boldsymbol{y}}_{\boldsymbol{t}})p(\tilde{\boldsymbol{y}}_{\boldsymbol{t}}|I_{t-1}, M)d\boldsymbol{v}_t d\tilde{\boldsymbol{y}}_{\boldsymbol{t}}, \qquad (2)$$

where I_{t-1} is the joint information set of all models and M the union of all models. The integrals are of dimension n and $n+1$ for each time observation.

Structure of our FDC: Choice of the different densities

A key step is to give content to the different densities.

- $p(y_t|\boldsymbol{v}_t, \tilde{\boldsymbol{y}}_t)$ is labeled the **multivariate normal combination density** :

$$p(y_t|\boldsymbol{v}_t, \tilde{\boldsymbol{y}}_t) = n(y_t|v_{0t} + \sum_{i=1}^{n} v_{it}\tilde{y}_{it}, \sigma_t^2), \tag{3}$$

  where time-varying constant v_{0t} in the conditional mean allows for forecast adjustments to shocks and regime changes in the data. σ_t^2 allows for time-varying volatility.

Structure of our FDC: Choice of the different densities

A key step is to give content to the different densities.

- $p(y_t|\boldsymbol{v}_t, \tilde{\boldsymbol{y}}_t)$ is labeled the **multivariate normal combination density** :

$$p(y_t|\boldsymbol{v}_t, \tilde{\boldsymbol{y}}_t) = n(y_t|v_{0t} + \sum_{i=1}^{n} v_{it}\tilde{y}_{it}, \sigma_t^2), \qquad (3)$$

  where time-varying constant v_{0t} in the conditional mean allows for forecast adjustments to shocks and regime changes in the data. σ_t^2 allows for time-varying volatility.

- $p(\boldsymbol{v}_t|\tilde{\boldsymbol{y}}_t)$ is labeled the **density of latent time-varying parameter weights** and specified as:

$$p(\boldsymbol{v}_t|\boldsymbol{v}_{t-1}, \boldsymbol{\Sigma}_t) = n(\boldsymbol{v}_t|\boldsymbol{v}_{t-1}, \boldsymbol{\Sigma}_t), \qquad (4)$$

  where the parameter $\boldsymbol{\Sigma}_t = \sigma_t^2 \boldsymbol{W}_t$ and $\boldsymbol{W}_t$ is a diagonal matrix with elements w_{it} given in the paper.

Structure of our FDC: Choice of the different densities

A key step is to give content to the different densities.

- $p(y_t|\boldsymbol{v}_t, \tilde{\boldsymbol{y}}_t)$ is labeled the **multivariate normal combination density** :

$$p(y_t|\boldsymbol{v}_t, \tilde{\boldsymbol{y}}_t) = n(y_t|v_{0t} + \sum_{i=1}^{n} v_{it}\tilde{y}_{it}, \sigma_t^2), \tag{3}$$

  where time-varying constant v_{0t} in the conditional mean allows for forecast adjustments to shocks and regime changes in the data. σ_t^2 allows for time-varying volatility.

- $p(\boldsymbol{v}_t|\tilde{\boldsymbol{y}}_t)$ is labeled the **density of latent time-varying parameter weights** and specified as:

$$p(\boldsymbol{v}_t|\boldsymbol{v}_{t-1}, \boldsymbol{\Sigma_t}) = n(\boldsymbol{v}_t|\boldsymbol{v}_{t-1}, \boldsymbol{\Sigma_t}), \tag{4}$$

  where the parameter $\boldsymbol{\Sigma_t} = \sigma_t^2 \boldsymbol{W_t}$ and $\boldsymbol{W_t}$ is a diagonal matrix with elements w_{it} given in the paper.

- $p(\tilde{\boldsymbol{y}}_t|I_{t-1}, M)$ is labeled **the joint forecast density of the different models.** Due to the conditional independence assumption it is given as:

$$p(\tilde{\boldsymbol{y}}_t|I_{t-1}, M) = \prod_{i=1}^{n} p(\tilde{y}_{it}|I_{i(t-1)}, M_i). \tag{5}$$

Structure of our FDC: The corresponding equation system is an econometric interpretation of Bayesian Predictive Synthesis approach from Mike West et al

- **The Equation System:** a **multivariate regression model** with **generated regressors** $\tilde{y}_t$, given as draws from the forecast distributions of the different models and **time-varying parameters** v_{it} draws:

$$y_t = v_{0t} + \sum_{i=1}^{n} v_{it} \tilde{y}_{it} + \varepsilon_t :: \varepsilon_t \sim NID(0, \sigma_t^2), t = 1, \ldots, T. \quad (6)$$

Structure of our FDC: The corresponding equation system is an econometric interpretation of Bayesian Predictive Synthesis approach from Mike West et al

- **The Equation System:** a **multivariate regression model** with **generated regressors** $\tilde{y}_t$, given as draws from the forecast distributions of the different models and **time-varying parameters** v_{it} draws:

$$y_t = v_{0t} + \sum_{i=1}^{n} v_{it}\tilde{y}_{it} + \varepsilon_t :: \varepsilon_t \sim NID(0, \sigma_t^2), t = 1, \ldots, T. \quad (6)$$

- where the latent time-varying parameters are specified to follow a **Random Walk learning process**:

$$v_{it} = v_{it-1} + \varepsilon_{vt} :: \varepsilon_{vt} \sim NID(0, \sigma_{vt}^2 = \sigma_t^2 w_t), i = 0, \ldots, n. \quad (7)$$

Structure of our FDC: The corresponding equation system is an econometric interpretation of Bayesian Predictive Synthesis approach from Mike West et al

- **The Equation System:** a **multivariate regression model** with **generated regressors** $\tilde{y}_t$, given as draws from the forecast distributions of the different models and **time-varying parameters** v_{it} draws:

$$y_t = v_{0t} + \sum_{i=1}^{n} v_{it}\tilde{y}_{it} + \varepsilon_t :: \varepsilon_t \sim NID(0, \sigma_t^2), t = 1, \ldots, T. \quad (6)$$

- where the latent time-varying parameters are specified to follow a **Random Walk learning process**:

$$v_{it} = v_{it-1} + \varepsilon_{vt} :: \varepsilon_{vt} \sim NID(0, \sigma_{vt}^2 = \sigma_t^2 w_t), i = 0, \ldots, n. \quad (7)$$

- where σ_{vt}^2 is defined via a standard single discount factor specification (see Prado and West (2010)) and σ_t^2 is the residual variance in predicting y_t based on past information and the set of individual forecast distributions.It follows a beta-gamma volatility model (also based on discounting)

Structure of our FDC: The corresponding equation system is an econometric interpretation of Bayesian Predictive Synthesis approach from Mike West et al

- **The Equation System:** a **multivariate regression model** with **generated regressors** $\tilde{y}_t$, given as draws from the forecast distributions of the different models and **time-varying parameters** v_{it} draws:

$$y_t = v_{0t} + \sum_{i=1}^{n} v_{it}\tilde{y}_{it} + \varepsilon_t :: \varepsilon_t \sim NID(0, \sigma_t^2), t = 1, \ldots, T. \quad (6)$$

- where the latent time-varying parameters are specified to follow a **Random Walk learning process**:

$$v_{it} = v_{it-1} + \varepsilon_{vt} :: \varepsilon_{vt} \sim NID(0, \sigma_{vt}^2 = \sigma_t^2 w_t), i = 0, \ldots, n. \quad (7)$$

- where σ_{vt}^2 is defined via a standard single discount factor specification (see Prado and West (2010)) and σ_t^2 is the residual variance in predicting y_t based on past information and the set of individual forecast distributions.It follows a beta-gamma volatility model (also based on discounting)
- Mike West et al calls this a DLM (Dynamic Linear Factor Model).

Learning from errors: Forecast errors and model set incompleteness

- The disturbance ε_t implied by the combination density is given as:

$$\varepsilon_t = y_t - (v_{0t} + \sum_{i=1}^{n} v_{it}\tilde{y}_{it}). \tag{8}$$

It is a **weighted** combination of forecast errors: $y_t - \tilde{y}_{it}, i = \ldots n$.
Forecast errors are due to:

  - **Sudden shocks in the series, volatility**
  - **Misspecification errors** from **model set incompleteness**

Learning from errors: Forecast errors and model set incompleteness

- The disturbance ε_t implied by the combination density is given as:

$$\varepsilon_t = y_t - (v_{0t} + \sum_{i=1}^{n} v_{it}\tilde{y}_{it}). \tag{8}$$

It is a **weighted** combination of forecast errors: $y_t - \tilde{y}_{it}, i = \ldots n$.
Forecast errors are due to:

  - **Sudden shocks in the series, volatility**
  - **Misspecification errors** from **model set incompleteness**

- The dynamic behaviour of the individual disturbance ε_{it} from model M_i given as:

$$\varepsilon_{it} = y_t - (v_{0,it} + v_{it}\tilde{y}_{it}), \tag{9}$$

which indicates the **weighted** forecast error in the i-the model.

Road Map of the Probability model as Generalised Linear State Space System

Time series models

$M_i \ (i = 1, \ldots, n)$:

$\tilde{y}_{it} \sim p(\tilde{y}_{it} | I_{i(t-1)}, M_i)$

Stochastic volatility model:

$\varepsilon_t \sim NID(0, \sigma_t^2)$

$\sigma_t^2 = \frac{\delta \sigma_{t-1}^2}{\gamma_t}$

$\gamma_t \sim \text{Beta}(\frac{\delta h_{t-1}}{2}, \frac{(1-\delta) h_{t-1}}{2})$

$h_t = \delta h_{t-1} + 1$

Central equation:

$y_t = v_{0t} + \sum_{i=1}^n v_{it} \tilde{y}_{it} + \varepsilon_t$

Random walk learning for

unrestricted latent variables:

$v_{it} = v_{i(t-1} + \varepsilon_{vt}$

$\varepsilon_{vt} \overset{iid}{\sim} \mathcal{N}(0, \sigma_{vt}^2 = \sigma_t^2 w_t)$

$w_t = \frac{\beta - 1}{\beta} w_{t-1}$

Outline of Algorithm

3 stage Markov Chain Monte Carlo

1. **Forecast from n models.** Generate draws from the forecast distributions from the n different models which gives $\tilde{y}_{it}, i = 1, \ldots, n$

Outline of Algorithm

3 stage Markov Chain Monte Carlo

1. **Forecast from n models.** Generate draws from the forecast distributions from the n different models which gives $\tilde{y}_{it}, i = 1, \ldots, n$
2. **Latent variable parameters**. Using the **Kalman/Normal Filter method (which includes updating)** with initial value $v_{i0}, i = 1, \ldots, n$, generate variable parameters $v_{it}, i = 1, \ldots, n$ from the RW process.

Outline of Algorithm

3 stage Markov Chain Monte Carlo

1. **Forecast from n models.** Generate draws from the forecast distributions from the n different models which gives $\tilde{y}_{it}, i = 1, \ldots, n$
2. **Latent variable parameters**. Using the **Kalman/Normal Filter method (which includes updating)** with initial value $v_{i0}, i = 1, \ldots, n$, generate variable parameters $v_{it}, i = 1, \ldots, n$ from the RW process.
3. **SV parameters.** Given draws $\tilde{y}_{it}, i = 1, \ldots, n$, $v_t, i = 1, \ldots, n$, generate draw of the SV parameters from inverted Gamma distribution.

Outline of Algorithm

3 stage Markov Chain Monte Carlo

1. **Forecast from n models.** Generate draws from the forecast distributions from the n different models which gives $\tilde{y}_{it}, i = 1, \ldots, n$
2. **Latent variable parameters**. Using the **Kalman/Normal Filter method (which includes updating)** with initial value $v_{i0}, i = 1, \ldots, n$, generate variable parameters $v_{it}, i = 1, \ldots, n$ from the RW process.
3. **SV parameters.** Given draws $\tilde{y}_{it}, i = 1, \ldots, n$, $v_t, i = 1, \ldots, n$, generate draw of the SV parameters from inverted Gamma distribution.

1. **Forecasting proceeds as follows:**

Outline of Algorithm

3 stage Markov Chain Monte Carlo

1. **Forecast from *n* models.** Generate draws from the forecast distributions from the n different models which gives $\tilde{y}_{it}, i = 1, \ldots, n$
2. **Latent variable parameters**. Using the **Kalman/Normal Filter method (which includes updating)** with initial value $v_{i0}, i = 1, \ldots, n$, generate variable parameters $v_{it}, i = 1, \ldots, n$ from the RW process.
3. **SV parameters.** Given draws $\tilde{y}_{it}, i = 1, \ldots, n$, $v_t, i = 1, \ldots, n$, generate draw of the SV parameters from inverted Gamma distribution.

1. **Forecasting proceeds as follows:**
2. Given a generated $v_{it}, i = 1, \ldots, n$, a generated SV value, a generated $\tilde{y}_{it}, i = 1, \ldots, n$ and using (6) generate a one step predicted value y_{t+1}.

Outline of Algorithm

3 stage Markov Chain Monte Carlo

1. **Forecast from n models.** Generate draws from the forecast distributions from the n different models which gives $\tilde{y}_{it}, i = 1, \ldots, n$
2. **Latent variable parameters**. Using the **Kalman/Normal Filter method (which includes updating)** with initial value $v_{i0}, i = 1, \ldots, n$, generate variable parameters $v_{it}, i = 1, \ldots, n$ from the RW process.
3. **SV parameters.** Given draws $\tilde{y}_{it}, i = 1, \ldots, n$, $v_t, i = 1, \ldots, n$, generate draw of the SV parameters from inverted Gamma distribution.

1. **Forecasting proceeds as follows:**
2. Given a generated $v_{it}, i = 1, \ldots, n$, a generated SV value, a generated $\tilde{y}_{it}, i = 1, \ldots, n$ and using (6) generate a one step predicted value y_{t+1}.
3. Repeating this process gives a synthetic sample of future values and a forecast density at time $t + 1$.

Outline of Algorithm

3 stage Markov Chain Monte Carlo

1. **Forecast from n models.** Generate draws from the forecast distributions from the n different models which gives $\tilde{y}_{it}, i = 1, \ldots, n$
2. **Latent variable parameters**. Using the **Kalman/Normal Filter method (which includes updating)** with initial value $v_{i0}, i = 1, \ldots, n$, generate variable parameters $v_{it}, i = 1, \ldots, n$ from the RW process.
3. **SV parameters.** Given draws $\tilde{y}_{it}, i = 1, \ldots, n$, $v_t, i = 1, \ldots, n$, generate draw of the SV parameters from inverted Gamma distribution.

1. **Forecasting proceeds as follows:**
2. Given a generated $v_{it}, i = 1, \ldots, n$, a generated SV value, a generated $\tilde{y}_{it}, i = 1, \ldots, n$ and using (6) generate a one step predicted value y_{t+1}.
3. Repeating this process gives a synthetic sample of future values and a forecast density at time $t+1$.
4. Very Important feature from this MCMC procedure: **The uncertainty in the generated forecasts from the different models is directly carried forward in the uncertainty of the combined forecast density**. In contrast, **frequentists methods use a two-step method and they suffer from the generated regressor problem.**

Individual models

General framework for constructing forecast densities from individual models

- General stochastic volatility model with Student's t-distributed errors given by

$$S_{t+h|t} - \hat{S}_{t+h|t} = \varepsilon_{t+h|t}, \quad \varepsilon_{t+h|t} \sim T(\mu, e^{h_{t+h|t}}, \nu), \tag{10}$$

$$h_{t+h|t} = \mu + \phi(h_{t+h-1|t} - \mu) + \zeta_{t+h|t}, \quad \zeta_{t+h|t} \sim NID(0, \omega^2), \tag{11}$$

in which $|\phi| < 1$ and $\hat{S}_{t+h|t}$ is a point forecast of the real price.

General framework for constructing forecast densities from individual models

- General stochastic volatility model with Student's t-distributed errors given by

$$S_{t+h|t} - \hat{S}_{t+h|t} = \varepsilon_{t+h|t}, \quad \varepsilon_{t+h|t} \sim T(\mu, e^{h_t+h|t}, \nu), \tag{10}$$
$$h_{t+h|t} = \mu + \phi(h_{t+h-1|t} - \mu) + \zeta_{t+h|t}, \quad \zeta_{t+h|t} \sim NID(0, \omega^2), \tag{11}$$

in which $|\phi| < 1$ and $\hat{S}_{t+h|t}$ is a point forecast of the real price.

- Obtain draws from the forecast distribution of $\tilde{S}_{t+h|t}$, conditional on the model estimates

$$\tilde{S}_{t+h|t} = \hat{S}_{t+h|t} + \hat{\varepsilon}_{t+h|t}, \quad \varepsilon_t \sim T(0, e^{\hat{h}_{t+h|t}}, \hat{\nu}), \tag{12}$$

in which $\hat{\varepsilon}_{t+h|t}, \hat{h}_{t+h|t}$ and $\hat{\nu}$ are posterior draws from the estimated stochastic volatility model.

Individual forecasting models

- **No-change model (NC)**

$$\hat{S}_{t+h|t} = S_t. \tag{13}$$

Individual forecasting models

- **No-change model (NC)**

$$\hat{S}_{t+h|t} = S_t. \tag{13}$$

- **Changes in the price index of non-oil industrial raw materials (CRB)**

$$\hat{S}_{t+h|t} = S_{t|t}(1 + \pi_t^{h,rm} - \mathbb{E}_t[\pi_{t+h}^{(h)}]). \tag{14}$$

Individual forecasting models

- **No-change model (NC)**

$$\hat{S}_{t+h|t} = S_t. \tag{13}$$

- **Changes in the price index of non-oil industrial raw materials (CRB)**

$$\hat{S}_{t+h|t} = S_{t|t}(1 + \pi_t^{h,rm} - \mathbb{E}_t[\pi_{t+h}^{(h)}]). \tag{14}$$

- **Futures & West Texas Intermediate (WTI) oil futures prices (Futures)**

$$\hat{S}_{t+h|t} = S_{t|t}(1 + f_t^{WTI,h} - s_t^{\text{WTI}} - \mathbb{E}_t[\pi_{t+h}^{(h)}]), \tag{15}$$

Individual forecasting models

- **Spread & Spread Between the Spot Prices of Gasoline and Crude Oil (Spread)**

$$\hat{S}_{t+h|t} = S_{t|t}\exp(\hat{\beta}[s_t^{\text{gas}} - s_t^{\text{WTI}}] - \mathbb{E}_t[\pi_{t+h}^{(h)}]), \tag{16}$$

Individual forecasting models

- **Spread & Spread Between the Spot Prices of Gasoline and Crude Oil (Spread)**

$$\hat{S}_{t+h|t} = S_{t|t}\exp(\hat{\beta}[s_t^{\text{gas}} - s_t^{\text{WTI}}] - \mathbb{E}_t[\pi_{t+h}^{(h)}]), \quad (16)$$

- **Time-Varying Parameter Model of the Gasoline and Heating Oil Spreads (TVspread)**

$$\hat{S}_{t+h|t} = S_{t|t}\exp(\hat{\beta}_{1,t}[s_t^{\text{gas}} - s_t^{\text{WTI}}] + \hat{\beta}_{2,t}[s_t^{\text{heat}} - s_t^{\text{WTI}}] - \mathbb{E}_t[\pi_{t+h}^{(h)}]), \quad (17)$$

Individual forecasting models

- **Spread & Spread Between the Spot Prices of Gasoline and Crude Oil (Spread)**

$$\hat{S}_{t+h|t} = S_{t|t} \exp(\hat{\beta}[s_t^{\text{gas}} - s_t^{\text{WTI}}] - \mathbb{E}_t[\pi_{t+h}^{(h)}]), \tag{16}$$

- **Time-Varying Parameter Model of the Gasoline and Heating Oil Spreads (TVspread)**

$$\hat{S}_{t+h|t} = S_{t|t} \exp(\hat{\beta}_{1,t}[s_t^{\text{gas}} - s_t^{\text{WTI}}] + \hat{\beta}_{2,t}[s_t^{\text{heat}} - s_t^{\text{WTI}}] - \mathbb{E}_t[\pi_{t+h}^{(h)}]), \tag{17}$$

- **Oil market Vector Autoregression (VAR)**

$$y_t = b + \sum_{i=1}^{p} B_i y_{t-i} + e_t, \tag{18}$$

Empirical contributions

Empirical work

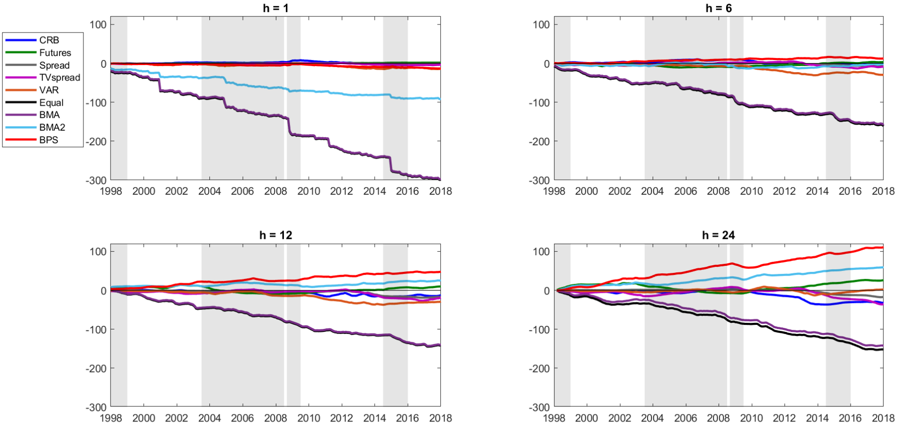
- Forecast monthly real price of crude oil
  - Real-time data as in Baumeister and Kilian (2012, 2015)
  - Training sample: 1992:01-1998:02
  - Evaluation sample: 1998:03-2017:12
  - Forecast evaluation: Root Mean Squared Forecast Error (RMSFE), Log Predictive Score (LPS) and their time behaviour, Time behaviour of weights and diagnostic measures.
  - Forecast horizons: $h = 1$, $h = 6$, $h = 12$, $h = 24$
- Consider different model combinations
  - BPS, BMA, BMA with rolling window weights, and equal weights

Density and point forecast results relative to a no-change benchmark, Evaluation sample 1998:03-2017:12

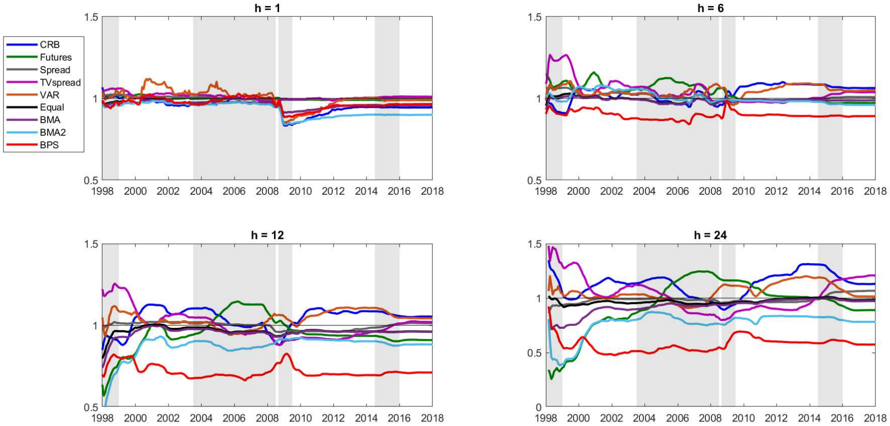
				LPS					
Horizon	CRB	Futures	Spread	TVspread	VAR	Equal	BMA	BMA2	BPS
1	0,46	**1,69***	-0,55	-3,62	-14,26	-298,95**	-297,04**	-91,44**	-13,23
6	-0,21	3,11	-5,10*	-8,14	-29,57**	-161,18**	-158,22**	-2,55*	**12,59****
12	-12,50	10,02*	-16,56**	-19,14**	-28,74**	-141,46**	-139,63**	25,85*	**47,73****
24	-32,48**	26,10**	-16,91**	-35,25**	2,34	-152,15**	-142,21**	59,14**	**110,96****

				RMSFE					
Horizon	CRB	Futures	Spread	TVspread	VAR	Equal	BMA	BMA2	BPS
1	0,95	0,99*	1,00	1,01	0,99	0,96*	0,96*	**0,90***	0,97
6	1,06*	0,97*	1,01	1,04	1,05*	0,99	0,99	0,96**	**0,89****
12	1,05	0,91**	1,01	1,02	1,04	0,96**	0,96**	0,88**	**0,71****
24	1,13**	0,89**	1,07	1,21**	1,01**	0,98	0,97**	0,78**	**0,57****

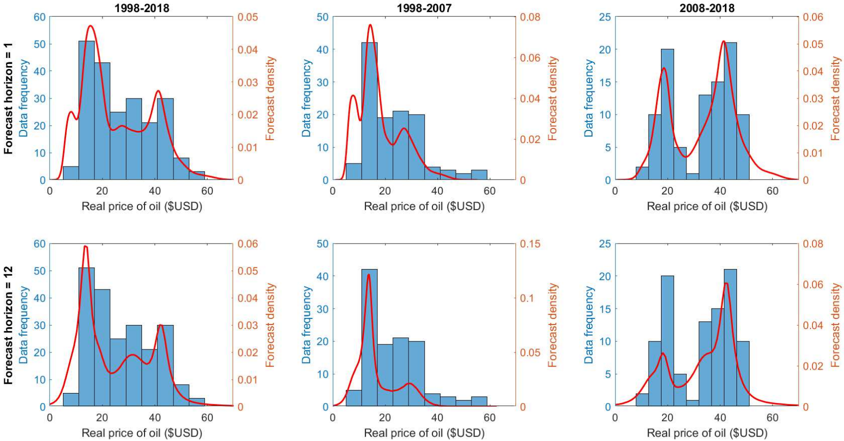
Time patterns of forecast means of cumulative Log Predictive Scores relative to a no-change model benchmark



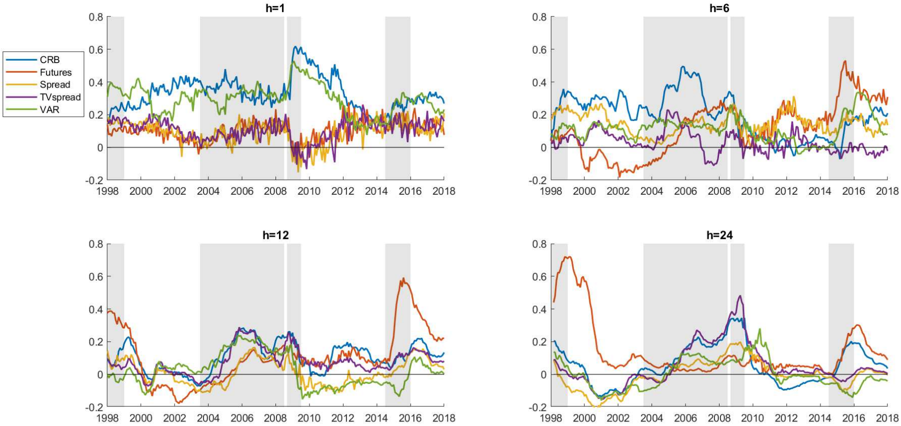
Time patterns of forecast means of Root Mean Squared Forecast Errors relative to a no-change model benchmark



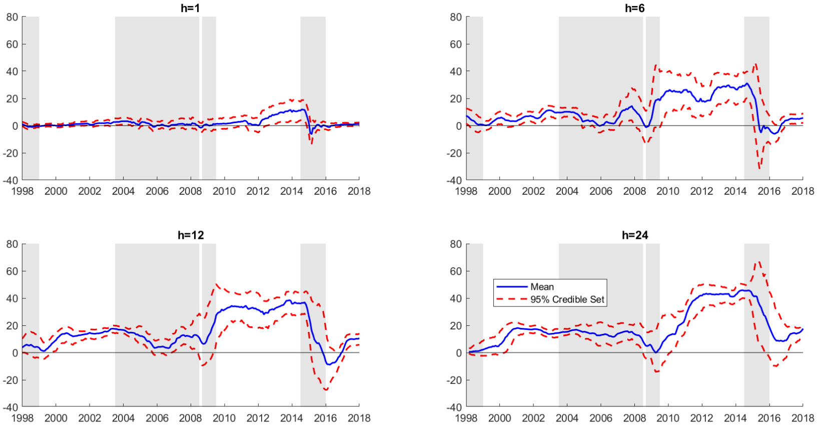
Observed data densities and estimated forecast density combinations pooled over specific subperiods



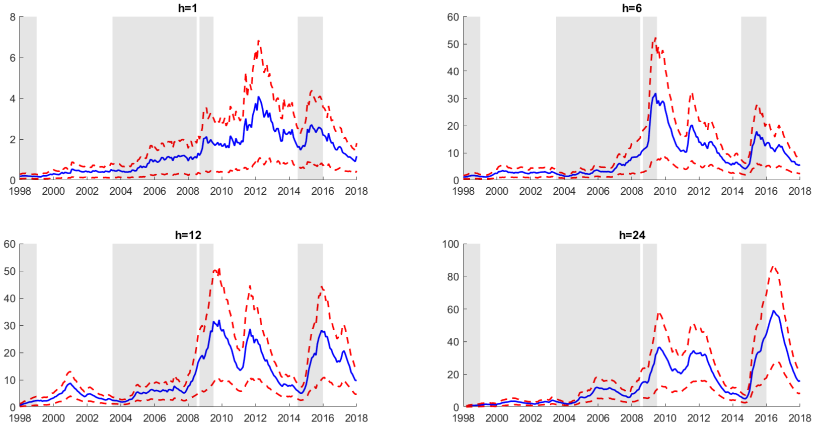
Time patterns of forecast means of model weights (v_{it}) in the FDC model based on BPS



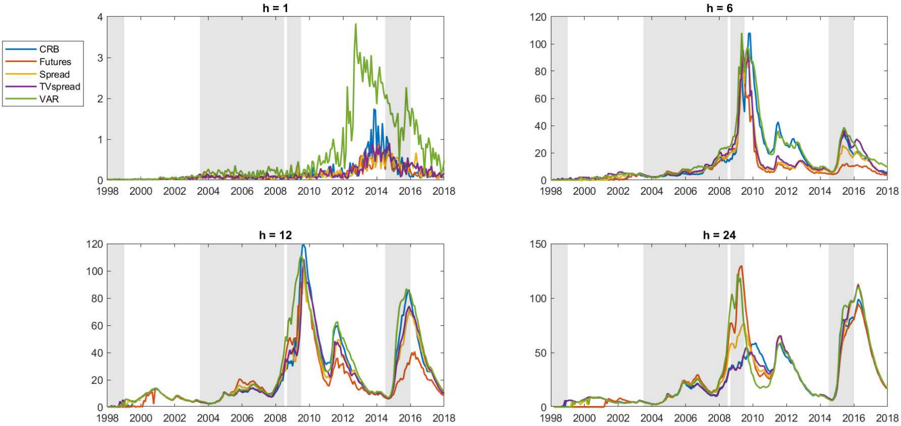
Time pattern of forecast means of intercept (v_{0t}) in the FDC model based on BPS



Time pattern of forecast means of variance (σ_t^2) for the central equation in BPS model.



Time patterns of forecast means of variances (σ^2_{it}) for individual models in the central equation in BPS model



Robustness checks

- Robustness check: **Alternative oil price series** We focused on forecasting the IRAC price of crude oil, which is commonly viewed as a proxy for the global price of oil. Two alternative series that are frequently cited in the press are the Brent and West Texas Intermediate (WTI) prices of crude oil.

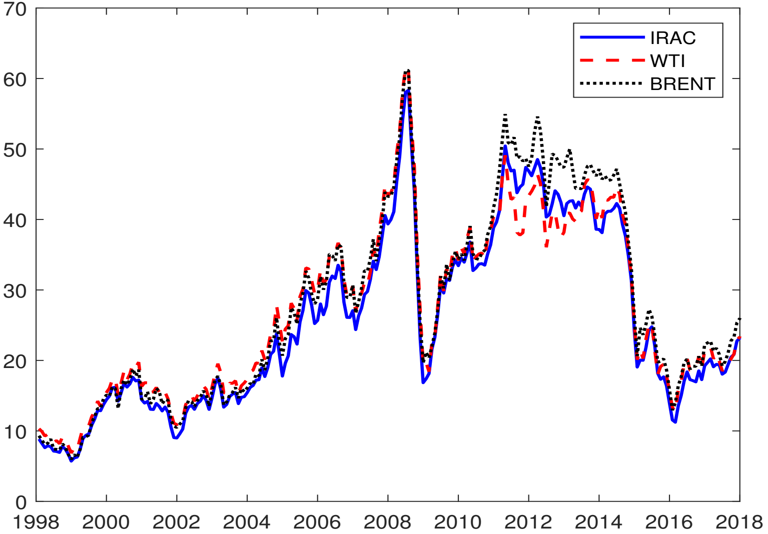
Robustness checks

- Robustness check: **Alternative oil price series** We focused on forecasting the IRAC price of crude oil, which is commonly viewed as a proxy for the global price of oil. Two alternative series that are frequently cited in the press are the Brent and West Texas Intermediate (WTI) prices of crude oil.
- Robustness check: **Alternative model sets** We estimated alternative models, see Table earlier. The main insight is that none of the additional models increase the forecast accuracy of the main BPS specification.

Robustness checks

- Robustness check: **Alternative oil price series** We focused on forecasting the IRAC price of crude oil, which is commonly viewed as a proxy for the global price of oil. Two alternative series that are frequently cited in the press are the Brent and West Texas Intermediate (WTI) prices of crude oil.
- Robustness check: **Alternative model sets** We estimated alternative models, see Table earlier. The main insight is that none of the additional models increase the forecast accuracy of the main BPS specification.
- Robustness check: **Alternative weighting procedures** Are time-varying parameter models as important as allowing for time-varying combination weights for our data set? In our data set time-varying combination weights in the main BPS specification are more important (give more accuracy) than individual time-varying parameter models.

Robustness checks: Alternative oil price series



Summary of results

1. Our combination approach systematically outperforms all benchmarks we compare it to
   - Gains in relative forecast accuracy are particularly substantial for **density forecast and at longer horizons**

Summary of results

1. Our combination approach systematically outperforms all benchmarks we compare it to
   - Gains in relative forecast accuracy are particularly substantial for **density forecast and at longer horizons**
2. The favourable forecast performance from our combination approach is not specific to certain time periods
   - **Large time variation** in the relative performance of the various individual models

Summary of results

1. Our combination approach systematically outperforms all benchmarks we compare it to
   - Gains in relative forecast accuracy are particularly substantial for **density forecast and at longer horizons**
2. The favourable forecast performance from our combination approach is not specific to certain time periods
   - **Large time variation** in the relative performance of the various individual models
3. Considerable time variation in the weights attached to each model
   - **Weights are not restricted to be a convex combination in the unit interval and can be negative.**

Summary of results

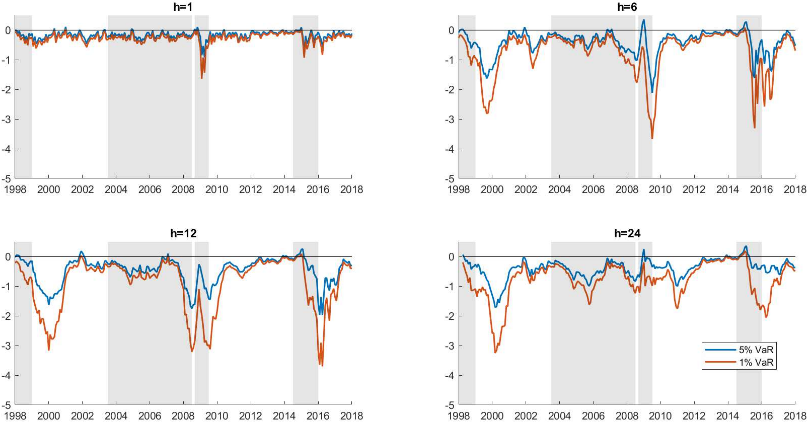
1. Our combination approach systematically outperforms all benchmarks we compare it to
   - Gains in relative forecast accuracy are particularly substantial for **density forecast and at longer horizons**
2. The favourable forecast performance from our combination approach is not specific to certain time periods
   - **Large time variation** in the relative performance of the various individual models
3. Considerable time variation in the weights attached to each model
   - **Weights are not restricted to be a convex combination in the unit interval and can be negative.**
4. Our combination is robust to model set incompleteness and misspecification
   - Time-varying intercept component that can adapt during episodes of low frequency signals
   - Built-in diagnostic information measures about forecast inaccuracy and/or model set incompleteness

Risk analysis

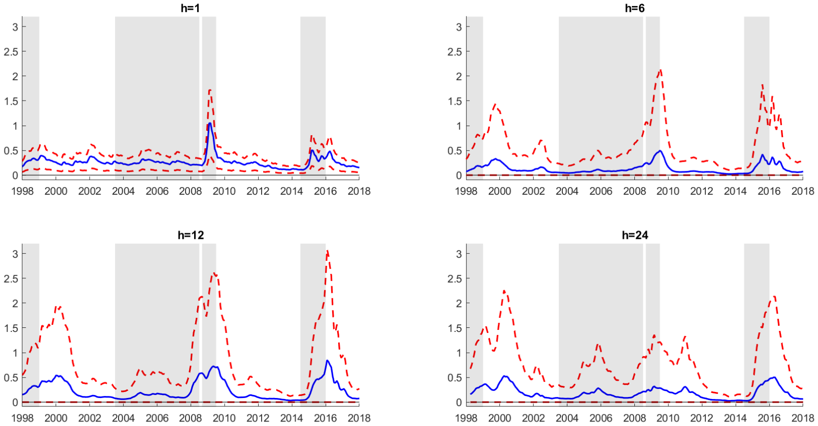
- Analyze the **risk and return** properties of investing in the global market for oil using our BPS modelling approach as an investment tool
  - Profit and loss distribution
  - Value-at-Risk (VaR)
  - Minimum Variance Hedge (MVH)

Profit and loss distribution (profit positive and loss negative) over the forecast evaluation period

Value at Risk for BPS over the forecast evaluation period



Minimum Variance Hedge ratios over the forecast evaluation period



Conclusion

- We have introduced a **basic probabilistic and numerically efficient Forecast Density Combination model**

Conclusion

- We have introduced a **basic probabilistic and numerically efficient Forecast Density Combination model**
- Our combination approach extends earlier approaches that have been applied to oil price forecasting models, by allowing for three key features
  1. **Time-varying and self-learning combination weights.**
  2. Explicitly modelling and estimation of time-varying forecast biases and facets of miscalibration of individual forecast densities and time-varying inter-dependencies among models
  3. Provide a **diagnostic learning analysis of model set incompleteness and learn from previous forecast mistakes.**

Conclusion

- We have introduced a **basic probabilistic and numerically efficient Forecast Density Combination model**
- Our combination approach extends earlier approaches that have been applied to oil price forecasting models, by allowing for three key features
  1. **Time-varying and self-learning combination weights.**
  2. Explicitly modelling and estimation of time-varying forecast biases and facets of miscalibration of individual forecast densities and time-varying inter-dependencies among models
  3. Provide a **diagnostic learning analysis of model set incompleteness and learn from previous forecast mistakes.**
- We have provided an extensive set of empirical results about time-varying forecast uncertainty and risk for the real price of oil

Topics for Further Research

- **Modelling and Data**. Use **more information** using many micro-data and explore diagnostic analysis about model incompleteness and richer model set.

Topics for Further Research

- **Modelling and Data**. Use **more information** using many micro-data and explore diagnostic analysis about model incompleteness and richer model set.
- More on **efficiency of filter** methods: Different filter methods like in machine learning: neural networks.

Topics for Further Research

- **Modelling and Data**. Use **more information** using many micro-data and explore diagnostic analysis about model incompleteness and richer model set.
- More on **efficiency of filter** methods: Different filter methods like in machine learning: neural networks.
- **Policy**. Applications to **Policy Issues**: In the field of Finance using Decision Models. Related paper by me and co-authors in *Journal of Econometrics 2019* is on **Forecast Density Combinations of Dynamic Models and Data Driven Portfolio Strategies** based on the concept of momentum patterns in the financial return data using US industrial portfolios 1929-2015. **Challenge to do this for macro-models**.

Topics for Further Research

- **Modelling and Data**. Use **more information** using many micro-data and explore diagnostic analysis about model incompleteness and richer model set.
- More on **efficiency of filter** methods: Different filter methods like in machine learning: neural networks.
- **Policy**. Applications to **Policy Issues**: In the field of Finance using Decision Models. Related paper by me and co-authors in *Journal of Econometrics 2019* is on **Forecast Density Combinations of Dynamic Models and Data Driven Portfolio Strategies** based on the concept of momentum patterns in the financial return data using US industrial portfolios 1929-2015. **Challenge to do this for macro-models**.
- **More efficient parallel computing.**

References I

Aastveit, K., K. Gerdrup, A. Jore, and L. Thorsrud (2014). Nowcasting GDP in real time: A density combination approach. *Journal of Business and Economic Statistics 32*(1), 48–68.

Aastveit, K. A., F. Ravazzolo, and H. K. Van Dijk (2018). Combined density nowcasting in an uncertain economic environment. *Journal of Business & Economic Statistics 36*(1), 131–145.

Alquist, R. and L. Kilian (2010). What do we learn from the price of crude oil futures? *Journal of Applied econometrics 25*(4), 539–573.

Alquist, R., L. Kilian, and R. J. Vigfusson (2013). Forecasting the price of oil. In *Handbook of economic forecasting*, Volume 2, pp. 427–507. Elsevier.

Amisano, G. and R. Giacomini (2007). Comparing density foreasts via weighted likelihood ratio tests. *Journal of Business & Economic Statistics 25*(2), 177–190.

Baumeister, C. and J. D. Hamilton (2019). Structural interpretation of vector autoregressions with incomplete identification: Revisiting the role of oil supply and demand shocks. *American Economic Review 109*(5), 1873–1910.

References II

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

Baumeister, C. and L. Kilian (2015). Forecasting the real price of oil in a changing world: A forecast combination approach. *Journal of Business & Economic Statistics 33*(3), 338–351.

Baumeister, C., D. Korobilis, and T. K. Lee (2020). Energy markets and global economic conditions. *The Review of Economics and Statistics*.

Bernard, J.-T., L. Khalaf, M. Kichian, and C. Yelou (2018). Oil Price Forecasts For The Long Term: Expert Outlooks, Models, Or Both? *Macroeconomic Dynamics 22*(3), 581–599.

Billio, M., R. Casarin, F. Ravazzolo, and H. K. van Dijk (2013). Time-varying combinations of predictive densities using nonlinear filtering. *Journal of Econometrics 177*, 213–232.

Casarin, R., S. Grassi, F. Ravazzolo, and H. K. van Dijk (2020). A bayesian dynamic compositional model for large density combinations in finance. Technical report.

References III

Casarin, R., S. Grassi, F. Ravazzolo, and H. vanDijk (2015). Parallel sequential monte carlo for efficient density combination: The deco matlab toolbox. *Journal of Statistical Software (Online) 68*.

Del Negro, M., R. B. Hasegawa, and F. Schorfheide (2016). Dynamic prediction pools: An investigation of financial frictions and forecasting performance. *Journal of Econometrics 192*(2), 391–405.

Garratt, A., S. P. Vahey, and Y. Zhang (2019). Real-time forecast combinations for the oil price. *Journal of Applied Econometrics 34*(3), 456–462.

Geweke, J. and G. Amisano (2011). Optimal prediction pools. *Journal of Econometrics 164*(1), 130 – 141.

Geweke, J. and G. G. Amisano (2012). Prediction with misspecified models. *The American Economic Review 102*, 482–486.

Gneiting, T. and R. Ranjan (2013). Combining predictive distributions. *Electron. J. Statist. 7*, 1747–1782.

Hall, S. G. and J. Mitchell (2007). Combining density forecasts. *International Journal of Forecasting 23*(1), 1–13.

References IV

Hamilton, J. D. (1983). Oil and the macroeconomy since world war ii. *Journal of political economy 91*(2), 228–248.

Hamilton, J. D. (2009). Understanding Crude Oil Prices. *The Energy Journal 30*(2), 179–206.

Hoogerheide, L., R. Kleijn, R. Ravazzolo, H. K. van Dijk, and M. Verbeek (2010). Forecast Accuracy and Economic Gains from Bayesian Model Averaging using Time Varying Weights. *Journal of Forecasting 29*(1-2), 251–269.

Jore, A. S., J. Mitchell, and S. P. Vahey (2010). Combining forecast densities from vars with uncertain instabilities. *Journal of Applied Econometrics 25*(4), 621–634.

Kascha, C. and F. Ravazzolo (2010). Combining inflation density forecasts. *Journal of Forecasting 29*(1-2), 231–250.

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), 1053–69.

Koop, G. and D. Korobilis (2012). Forecasting Inflation Using Dynamic Model Averaging. *International Economic Review 53*(3), 867–886.

Manescu, C. B. and I. V. Robays (2016). Forecasting the Brent Oil Price: Addressing Time-Variation in Forecast Performance. Technical report.

McAlinn, K., K. A. Aastveit, J. Nakajima, and M. West (2020). Multivariate bayesian predictive synthesis in macroeconomic forecasting. *Journal of the American Statistical Association 115*(531), 1092–1110.

McAlinn, K. and M. West (2019). Dynamic bayesian predictive synthesis in time series forecasting. *Journal of Econometrics 210*(1), 155–169.

Pak, A. (2018). Predicting crude oil prices: Replication of the empirical results in what do we learn from the price of crude oil?. *Journal of Applied Econometrics 33*(1), 160–163.

Pettenuzzo, D. and F. Ravazzolo (2016). Optimal portfolio choice under decision-based model combinations. *Journal of Applied Econometrics 31*, 1312–1332.

Ravazzolo, F. and P. Rothman (2013). Oil and U.S. GDP: A Real-Time Out-of-Sample Examination. *Journal of Money, Credit and Banking 45*(2-3), 449–463.

Takanashi, K. and K. McAlinn (2020). Predictive properties and minimaxity of bayesian predictive synthesis.