Darmstadt
Discussion Papers
in ECONOMICS

TECHNISCHE UNIVERSITÄT DARMSTADT

Improving Oil Price Forecasts by Sparse VAR Methods

Jens J. Krüger and Sebastian Ruths Sion

Nr. 237

Arbeitspapiere der Volkswirtschaftlichen Fachgebiete der TU Darmstadt

ISSN: 1438-2733

ISSN 1438-2733

Printed in Germany
Technische Universität Darmstadt
Department of Law and Economics
D – 64289 Darmstadt
Germany
www.wi.tu-darmstadt.de

Improving Oil Price Forecasts by Sparse VAR Methods

December 2019

by

Jens J. Krüger and Sebastian Ruths Sion

Darmstadt University of Technology
Department of Law and Economics
Hochschulstr. 1, D-64289 Darmstadt, Germany
tel.: +49 6151 1657 281, fax: +49 6151 1657 282
e-mail: krueger@vwl.tu-darmstadt.de

Abstract

In this paper we document the results of a forecast evaluation exercise for the real world price of crude oil using VAR models estimated by sparse (regularization) estimators. These methods have the property to constrain single parameters to zero. We find that estimating VARs with three core variables (real price of oil, index of global real economic activity, change in global crude oil production) by the sparse methods is associated with substantial reductions of forecast errors. The transformation of the variables (taking logs or differences) is also crucial. Extending the VARs by further variables is not associated with additonal gains in forecast performance as is the application of impulse indicator saturation before the estimation.

JEL classification: C32, Q47

Keywords: oil price prediction, vector autoregression, regularization

1 Introduction

The oil price and its changes have been associated with U.S. macroeconomic aggregates as well as the global business cycle (see e.g. Hamilton (1983), Kilian (2008)). The important oil price shocks in the 1970s and 1980s gained widespread attention in the public. As of today, crude oil is indispensable for keeping standards of living in developed economies as well as for fueling economic growth in rapidly developing nations such as China and India. Therefore, knowledge about the future price of oil is of importance for different actors. Researchers in central banks and international organizations such as the IMF use oil price forecasts as input in their forward looking macroeconomic models (see Baumeister and Kilian (2014)). Thus, improving crude oil price forecasts helps generating better macroeconomic projections as well as better future risk assessment associated with oil price fluctuations. Oil price forecasts also are helpful for governments of oil exporting countries which strongly depend on oil revenues to finance public expenses in budget planning. Relatedly, forecasting the oil price aids governments of countries that heavily rely on crude oil imports in shaping their environmental policies and energy tax setting. Improved oil price forecasts also support firms in their investment and purchasing decisions. For example, airlines and automobile companies take oil price forecasts into consideration when they decide about fares and product prices as well as product portfolios. Similarly, private homeowners might upgrade to energy-saving heating systems when forecasts point to future heating oil price increases.

In this paper we start from the global three-variable vector autoregressive (VAR) model for crude oil as first proposed by Kilian (2009) as the benchmark and investigate variants with enhanced variable sets using sparse (regularization) methods in order to evaluate and compare their forecast properties. Sparse estimation methods gained widespread attention in the machine learning literature (see e.g. Murphy (2012)) and now find more and more economic applications as a variable selection procedure. This is particularly important for vector autoregressions where a large number of parameters are to be estimated and usually only a common lag length for all equations is selected by information criteria.

Also in previous research on oil price forecasting an increasing trend towards basing the forecasts on a broader information set can be observed. This strand of research is mostly focused on applications of forecast combination methods (see Baumeister et al. (2014), Baumeister and Kilian (2015), Funk (2018), Garratt et al. (2019), Wang et al. (2017), Zhang et al. (2018)) or model averaging (see Drachal (2016) and Naser (2016)). More recently neural networks as well as regularization methods also have been employed to improve oil price forecasts (see Chen et al. (2019), Zhang et al. (2019)).

The lessons from the forecast evaluation exercise reported in this paper can be summarized as follows. First, the results show that the lag order commonly fixed at 12 or 24 months, which is justified for impulse-response analysis, is detrimental to forecast performance. Second, appropriate variable transformation (logs, differences, levels) is crucial for the forecast performance. Third, applying sparse estimators leads to improvements in forecast performance when using the variable transformation originally employed by Kilian (2009) and in the VAR in levels. Regularization also improves forecasts for shorter horizons when we express the variables in differences. Finally, when augmenting the core variable set by industrial production indices, exchange rates and financial variables, regularization does not lead to forecast improvements and we even observe forecast deterioration in some occasions.

The paper unfolds as follows. Section 2 introduces the VAR framework as well as the three sparse estimation methods which are subsequently applied. Section 3 presents the core data series and discusses data transformation and stationarity assessment. In section 4 we discuss different basic VAR specifications and select the best performing one as the benchmark. We proceed by estimating the three-variable VAR with the sparse methods and evaluate the forecast performance in comparison to the selected benchmark. Section 5 extends the three core variables by further variable sets containing production indices, exchange rates, financial variables and impulse indicator saturation dummies and evaluates the forecast performance. We conclude in section 6 with the discussion of the main findings.

2 Sparse VAR Methods

Before we turn to the description of the data used and how we dealt with stationarity-integration issues we briefly explain the forecasting models used. Our forecast evaluation exercise relies on the framework of a vector autoregression (VAR), pioneered by Sims (1980). In addition, we explain the approaches to regularization which we employ to prune the parameter matrices to obtain a more parsimonious (sparse) model specification, more precise parameter estimates and possibly reduced forecast errors. It is the primary aim of this study to investigate the latter issue.

A vector autoregression is stated as a VAR(p) with p lags for m variables in the vector $\boldsymbol{y}_t = (y_{1t}, ..., y_{mt})'$ observed for the periods $t = 1, ..., T$,

$$\boldsymbol{y}_t = \boldsymbol{c} + \boldsymbol{A}_1 \boldsymbol{y}_{t-1} + ... + \boldsymbol{A}_p \boldsymbol{y}_{t-p} + \boldsymbol{u}_t. \tag{1}$$

A VAR can be consistently estimated by least squares equation by equation, which amounts to minimizing the sum of squared residuals

$$SSR(\boldsymbol{\theta}) = \sum_{t=p+1}^{T} \boldsymbol{u}_t' \boldsymbol{u}_t \tag{2}$$

as the objective function, where the parameter vector $\boldsymbol{\theta}$ is understood to stack all $k = m + pm^2$ parameters to be estimated (i.e. $\boldsymbol{c}$ and $\boldsymbol{A}_1, ..., \boldsymbol{A}_p$). See Lütkepohl (2005) as well as Kilian and Lütkepohl (2017) for a comprehensive account of VAR models.

Given the estimates for $\boldsymbol{c}$ and $\boldsymbol{A}_1, ..., \boldsymbol{A}_p$, denoted $\hat{\boldsymbol{c}}$ and $\hat{\boldsymbol{A}}_1, ..., \hat{\boldsymbol{A}}_p$, respectively, the VAR can be used for generating forecasts by iterating equation (1) forward. This leads to forecasts one step and two steps into the future, written as

$$\hat{\boldsymbol{y}}_{T+1|T} = \hat{\boldsymbol{c}} + \hat{\boldsymbol{A}}_1 \boldsymbol{y}_T + ... + \hat{\boldsymbol{A}}_p \boldsymbol{y}_{T-p+1} \tag{3}$$

and

$$\hat{\boldsymbol{y}}_{T+2|T} = \hat{\boldsymbol{c}} + \hat{\boldsymbol{A}}_1 \hat{\boldsymbol{y}}_{T+1|T} + ... + \hat{\boldsymbol{A}}_p \boldsymbol{y}_{T-p+2}, \tag{4}$$

respectively. Here $\hat{\boldsymbol{y}}_{T+1|T}$ denotes the forecast for the variables one time step into the future given that the available information ends in period T. Note that for the 2-step forecast $\hat{\boldsymbol{y}}_{T+2|T}$ the first lag on the right hand side would be $\boldsymbol{y}_{T+1}$ which is not available in the data (the sample ends in period T) and is therefore substituted by the 1-step forecast $\hat{\boldsymbol{y}}_{T+1|T}$. In general, the h-step forecasts generated by conditional expectations are estimates of the conditional expectation $E(\boldsymbol{y}_{T+h} \,|\, \boldsymbol{y}_T, ..., \boldsymbol{y}_1)$. The h-step forecasts are computed by

$$\hat{\boldsymbol{y}}_{T+h|T} = \hat{\boldsymbol{c}} + \hat{\boldsymbol{A}}_1 \hat{\boldsymbol{y}}_{T+h-1|T} + ... + \hat{\boldsymbol{A}}_p \hat{\boldsymbol{y}}_{T+h-p|T}, \tag{5}$$

upon the substitution $\hat{\boldsymbol{y}}_{T+j|T} = \boldsymbol{y}_{T+j}$ whenever $j \leq 0$. Forecasts constructed in this way minimize the theoretical mean squared error (MSE).

The number of parameters arising in unconstrained VAR with lag length p is usually quite large, i.e. $k = m + pm^2$. Not all those parameters are different from zero although their estimates are so by chance and this may be detrimental to forecast performance. Since information criteria for lag order selection only eliminate entire parameter matrices $\boldsymbol{A}_j$, it would be helpful to use statistical methods which constrain selective parameters within these matrices to be zero. In the statistical literature this is known as sparsity or regularization to reduce the number of parameters which are different from zero.

Typical methods for regularization are the LASSO, the Elastic Net and the SCAD method which are explained below. These methods have in common that a penalty term $P(\boldsymbol{\theta})$ for the magnitude of the parameters is added to the objective function to be minimized

$$Z(\boldsymbol{\theta}) = SSR(\boldsymbol{\theta}) + \lambda P(\boldsymbol{\theta}) \tag{6}$$

with the penalty weight $\lambda > 0$ to be determined by cross-validation techniques. Hastie et al. (2009) provide a lucid exposition of variable selection by regularization methods (also known as shrinkage methods) in general.

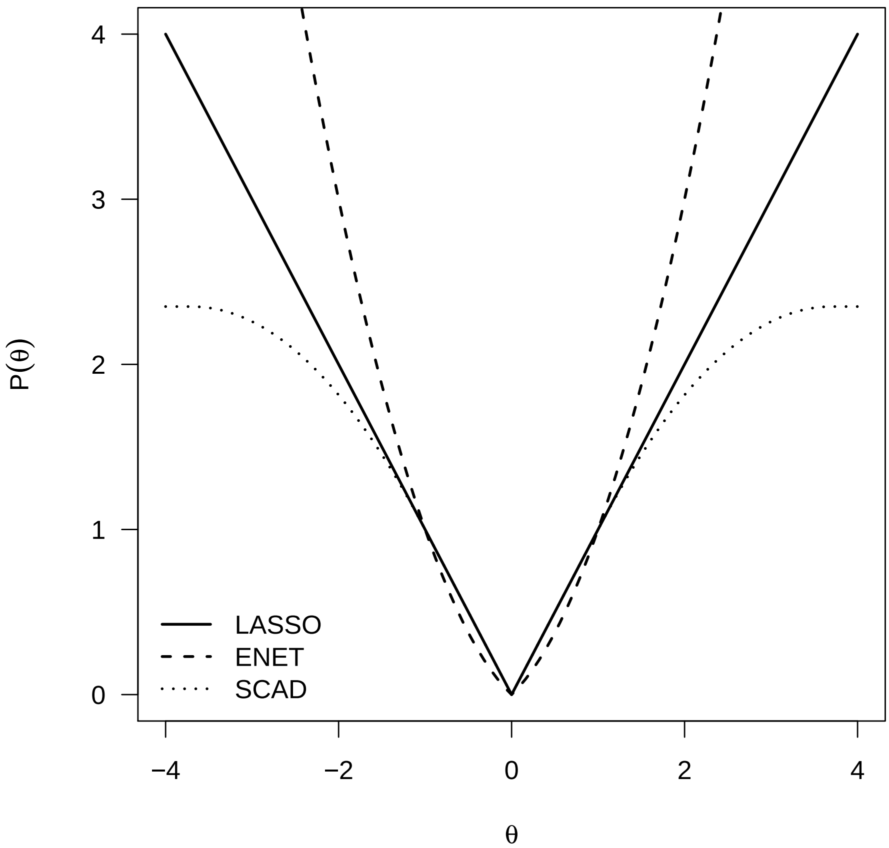
In this work we investigate the forecast performance of VARs estimated by three common variants of regularization methods. First, the LASSO (least absolute shrinkage and selection operator) by Tibshirani (1996) specifies the penalty term as $P(\boldsymbol{\theta}) = \sum_{j=1}^{k} |\theta_j|$.[1] This constrains some of the parameter estimates to be exactly equal to zero and thus eliminates some of the lags of the corresponding variables in the VAR to reach sparsity. Second, the Elastic Net (ENET) by Zou and Hastie (2005) chooses $P(\boldsymbol{\theta}) = \sum_{j=1}^{k} \left(\alpha|\theta_j| + (1-\alpha)\theta_j^2\right)$ which is the a combination of the LASSO and Ridge penalties with α usually fixed at 0.5. Third, SCAD (smoothly clipped absolute deviation) by Fan and Li (2001) is based on $P(\boldsymbol{\theta}) = \sum_{j=1}^{k} p(\theta_j)$ with $p(\theta_j) = |\theta_j|$ if $|\theta_j| \leq \lambda$, $p(\theta_j) = (2\gamma|\theta_j| - \theta_j^2/\lambda - \lambda)/2(\gamma - 1)$ if $\lambda < |\theta_j| \leq \gamma\lambda$ and $p(\theta_j) = \lambda(\gamma+1)/2$ if $|\theta_j| > \gamma\lambda$ with $\gamma > 2$ (setting $\gamma = 3.7$ is recommended by Fan and Li (2001) as providing “good practical performance for various variable selection problems”). The SCAD penalty coincides with the LASSO for $|\theta_j| \leq \lambda$, is a concave quadratic function until $|\theta_j| \leq \gamma\lambda$ and is constant for $|\theta_j| > \gamma\lambda$. This relaxes the intensity of penalization when the absolute value of the parameter increases.[2]

Figure 1 shows the penalty functions for the three sparse variants considered depicted for a scalar parameter θ (setting $\lambda = 1$, $\alpha = 0.5$ and $\gamma = 3.7$). All computations in this paper are performed using the packages “vars” and “sparsevar” for R.

[1] This is in contrast to Ridge regression, introduced by and Hoerl and Kennard (1970), using the penalty term $P(\boldsymbol{\theta}) = \sum_{j=1}^{k} \theta_j^2$ which serves to shrink the parameter estimates towards zero but does not set some of them exactly equal to zero as the LASSO does.

[2] A quite similar suggestion, called minimax concave penalty (MCP), has been made by Zhang (2010), which leads to results which are almost indistinguishable from SCAD and is therefore not further considered in the empirical forecast evaluation exercise of this paper.

Figure 1: Penalty Functions

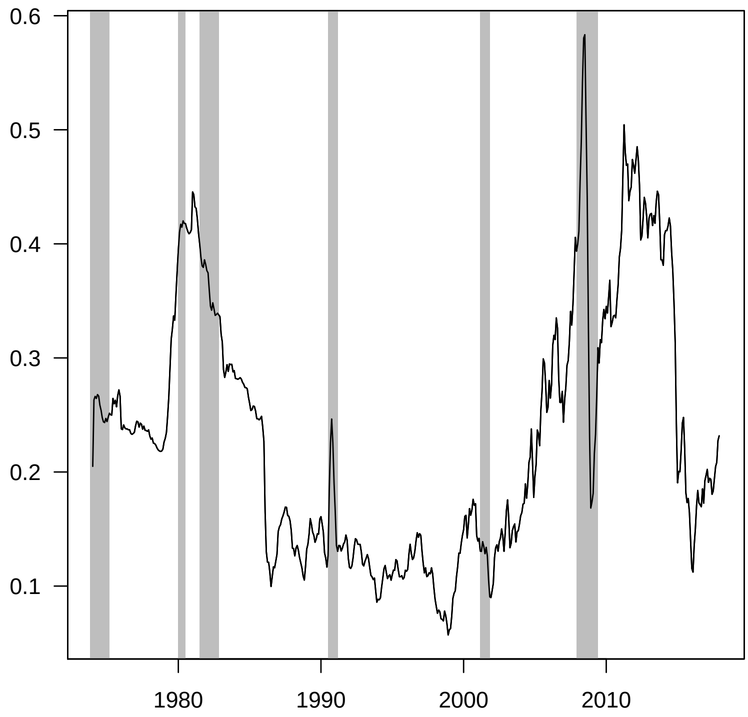


3 Data and Stationarity

In this section we discuss the properties of the three core variables used in the VAR models in section 4. These variables are used by Kilian (2009) in his structural VAR model to capture the main dynamics of the global market for crude oil as well as to estimate historical oil price shocks. The same variables are also used for the evaluation of oil price forecasting in the handbook article by Alquist, Kilian and Vigfusson (2013). Later, in section 5, these variables are extended by further sets of variables.

Before we turn to the core variables, their data sources and properties we have a closer look at the real price of crude oil, the target variable of our forecast evaluation exercise. Figure 2 shows the real oil price (in unlogged form) with NBER recession bars depicted in gray shades. This plot shows several distinct episodes of the history of the oil market over the last 50 years. The period starts with the years 1974-1986 which can be named the age of OPEC. This subperiod is characterized by the instability in the Middle East in the aftermath of the the of the Yom Kippur war, with the subsequent oil embargo, the Iranian revolution, the Iran-Iraq war and at least two major recessions in western countries. During these years we can also observe a worldwide monetary expansion leading to a depreciation of the US dollar and a higher oil price in dollars which compensated for the decline in revenues (purchasing power) of crude oil producers. The subperiod 1986-1997/8 witnessed the collapse of the OPEC cartel, an increase in worldwide oil supply and

Figure 2: Real Oil Price (1974-2017)

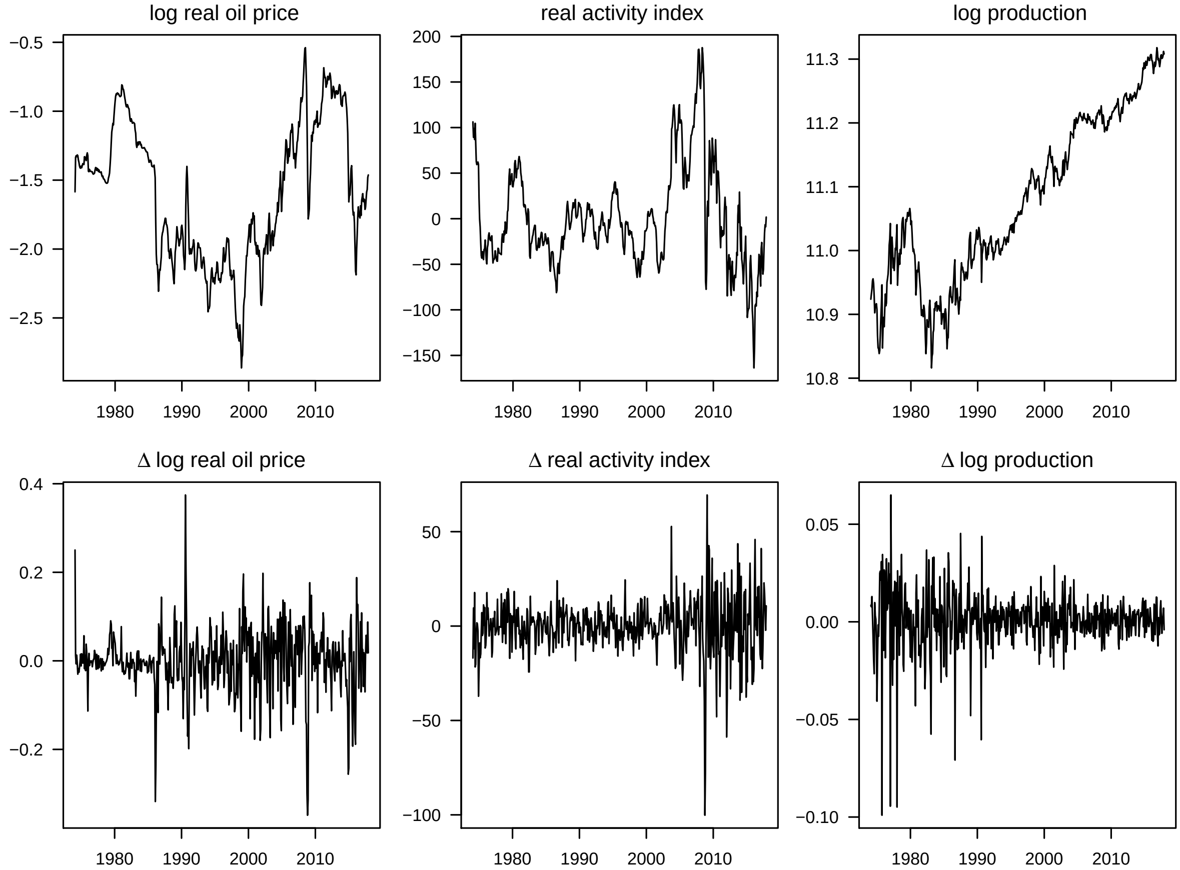


a lasting decline in world oil consumption, which kept the oil price low. The subperiod 1997/8-2008 was a phase of worldwide economic expansion, especially driven by emerging economies in east Asia such as China and India (see Hamilton (2009), Kilian and Hicks (2013)), with increasing demand for oil combined with slowing growth of crude oil supply. In the following the financial crisis of 2007/8 triggered the Great Recession in the US and lead to a prolonged worldwide economic downturn with the consequence of a collapse of world oil demand. Later in this period oil supply increased as a result of the fracking boom in the US, keeping the oil price under pressure. We refer to Baumeister and Kilian (2016) for a historical review of the oil price evolution and the main influencing episodes since 1973/74.

The real oil price is one of the three core variables used by Kilian (2009) and Alquist et al. (2013). More specifically, the three core variables are the real price of crude oil (deflated by the US consumer price index and in contrast to the discussed figure above expressed in logs) rpo_t, the index of global real economic activity as by Kilian (2009) rea_t and the percentage change in global crude oil production (computed as log differences) $\Delta prod_t$. See Kilian (2009) for a more thorough discussion of the construction of the variables, in particular regarding the real activity index. The index is constructed using dry cargo shipping rates based on the idea that global economic activity is the main driver of demand for international freight transport services. The updated data are retrieved from the homepage of Lutz Kilian (http://www-personal.umich.edu/~lkilian/) and incorporate the updates addressed in Kilian (2018) based on the methodological critique of Hamilton (2018).

The time series of the three core variables are plotted in figure 3. The sample period in this paper spans January 1974 to December 2017 implying a total sample size of $T = 528$

Figure 3: Time Series of the Core Variables



months. This extends the sample period of Kilian (2009), which goes until December 2007, now also comprising the time of the financial crisis, the breakdown of Lehman Brothers, the Great Recession and the recovery thereafter. For the real price of crude oil we use the refiners acquisition price of imported oil deflated by US CPI proposed by Kilian (2009) as the best measure for global oil prices.[3]

In the first panel the real oil price is now expressed in logs. The first row of the figure shows a trend in the production series and long swings of the real oil price and to a lesser degree in the case of the real activity index, pointing to a substantial degree of persistence. Both trend and persistence are characteristics of unit root nonstationarity (with and without a drift component, respectively). Therefore, this visual inspection suggests using the transformations (Δrpo_t, Δrea_t, $\Delta prod_t$) for the three core variables in the VAR.

When we try to confirm this by formal statistical testing the results (not shown in detail here) are mixed. For all three variables we find a strong rejection of the stationarity null hypothesis using the KPSS test of Kwiatkowski et al. (1992). This test is, however, prone to severe size distortions and therefore leads to substantial overrejections of the null hypothesis also under stationarity. The unit root null hypothesis is, however, also rejected in the case of the (log) real oil price using the DF-GLS test (or ERS test) of

[3]The refiners acquisition price of imported oil and global oil production series (in thousand barrels per day) are retrieved from the US Energy Information Administration. US CPI is retrieved from the FRED database with the series code CPIAUCSL.

Elliott et al. (1996). For the real activity index and (log) oil production, the unit root null can not be rejected. This is not overly surprising for the production series, but is somewhat puzzling in the case of the real activity index and its appearance in the figure 3. Applying the testing procedure to the first differences of the three variables we observe strong rejections of the unit root null, jointly with no rejections of the stationarity null. This is consistent with the visual inspection.

For control purposes and to achieve consistency with the literature (especially Kilian (2009) and Alquist et al. (2013)) we also perform the forecast evaluation exercise for the transformations (rpo_t, rea_t, $\Delta prod_t$), where, as defined above, the real price of oil and the production are expressed in logs. Regardless of the transformations applied, the target variable of the forecasts is the real price of oil (unlogged) which is the variable decisions makers are most likely to focus on rather than the corresponding logs or growth rates (differences of logs). According to Sims et al. (1990) determining the correct order of integration is not problematic for consistent parameter estimation in VAR models and should therefore not be problematic for forecasting.

4 Results with the Core Variables

In this section we discuss the results from the forecast evaluation exercise based on the three core variables. We first compare VARs with lag lengths fixed at $p = 12$ and $p = 24$, a VAR with lag length selected by AIC, and naïve no-change prediction (average real oil price over the previous 12 months) to select a benchmark for the sparse VAR methods. In the second step we evaluate the performance of the sparse VARs, estimated by LASSO, ENET and SCAD, in comparison with this benchmark model for the forecast horizons $h = 1, 2, 3, 6, 9, 12$.

Regarding the transformations of the three core variables we distinguish the Kilian VAR with variables $\boldsymbol{y}_t = (rpo_t,\, rea_t,\, \Delta prod_t)'$ as analyzed in Kilian (2009) and Alquist et al. (2013), the VAR in differences with variables $\boldsymbol{y}_t = (\Delta rpo_t,\, \Delta rea_t,\, \Delta prod_t)'$ and the VAR in levels with $\boldsymbol{y}_t = (rpo_t,\, rea_t,\, prod_t)'$. Using a VAR in levels, irrespective of the orders of integration of the variables and possible cointegration among the variables, is a standard approach in some fields, e.g. the empirical assessment of monetary policy (see e.g. Christiano et al. (2005)). There is a also considerable literature on the forecast performance of VARs in levels versus first differences (see e.g. Hoffman and Rasche (1996)).

When we suppose that all variables are integrated of order one and we are indeed able to establish cointegration by the Johansen (1988, 1991) trace test. Using an expanding test sample size starting from the first 100 observations up to the total sample we can establish cointegration for most of the samples before the financial crisis which is substantially weakened by the impact of the crisis. In the presence of cointegration the Granger representation theorem (Engle and Granger (1987)) justifies the estimation of a VAR in levels as a reduced form basis for forecasting. Even in the absence of cointegration there are good arguments that the decision between differences or levels is rather inessential when the VAR is used for forecasting. As explained by Kilian and Lütkepohl (2017, pp. 373f.) the main reason is the inherent ability of the VAR in levels to encompass a VAR model with integrated and possibly cointegrated variables as well as a VAR for stationary time series. This argument is reinforced by the uncertainty about unit root and cointegration properties of the time series and the often neglected fact that deciding between a VAR in differences and a cointegrating VAR is also subject to pre-testing bias.

Depending on the specific transformation of the real oil price, we obtain a forecast of the log (Kilian VAR and VAR in levels) or of the log differences (VAR in differences) of the real oil price variable. To compare these forecasts with the unlogged real oil price as our target variable, the forecasts are appropriately re-transformed (meaning taking exponentials when the real oil price has been logged or cumulating growth rates starting from the last observation in the data).

The forecast experiment is specified with an expanding window for the estimation sample with the first sample spanning 20 years (240 months) from January 1974 until December 1993 and the first forecast for January 1994 for a horizon $h = 1$ (February 1994 for $h = 2$, March 1994 for $h = 3$, June 1994 for $h = 6$, September 1994 for $h = 9$ and December 1994 for $h = 12$). Note that in the subsequent figures all forecast error measures are aligned at the position of the final observation of the estimation sample (i.e. December 1993 in the case of the first forecast) irrespective of the forecast horizon. Then the procedure is repeated with a further month, January 1994, added to the estimation sample. Proceeding in this way month by month we end up with a final estimation sample from January 1974 until December 2016 (43 years or 516 months) with forecasts for January 2017 ($h = 1$) until December 2017 ($h = 12$) which are all assigned to December 2016 in the figures.[4]

4.1 Benchmark VAR

The results for four candidates of our benchmark model are shown in figure 4. The curves show the recursive mean-squared error (MSE) measures[5] for the VAR(24) with a fixed lag length of $p = 24$ (VAR24, dotted line), used by Alquist et al. (2013), the VAR(12) with reduced lag length of $p = 12$ (VAR12, dashed line), VAR(AIC) with the lag length p chosen by the Akaike Information Criterion (VARAIC, dash-dotted line)[6] and the naïve no-change forecasts (solid line), which are used as the benchmark forecast in Alquist et al. (2013).

Each column pertains to a different transformation of the three variables (from left to right: VAR with transformation as in Kilian (2009), VAR in differences, VAR in levels) while the rows show the results for a particular forecast horizon of $h \in \{1, 2, 3, 6, 9, 12\}$ months. The horizontal lines indicate the smallest recursive MSE value at the end of the evaluation period which is achieved by any of the methods under consideration. The numerical value of this smallest MSE is printed directly above the horizontal line.

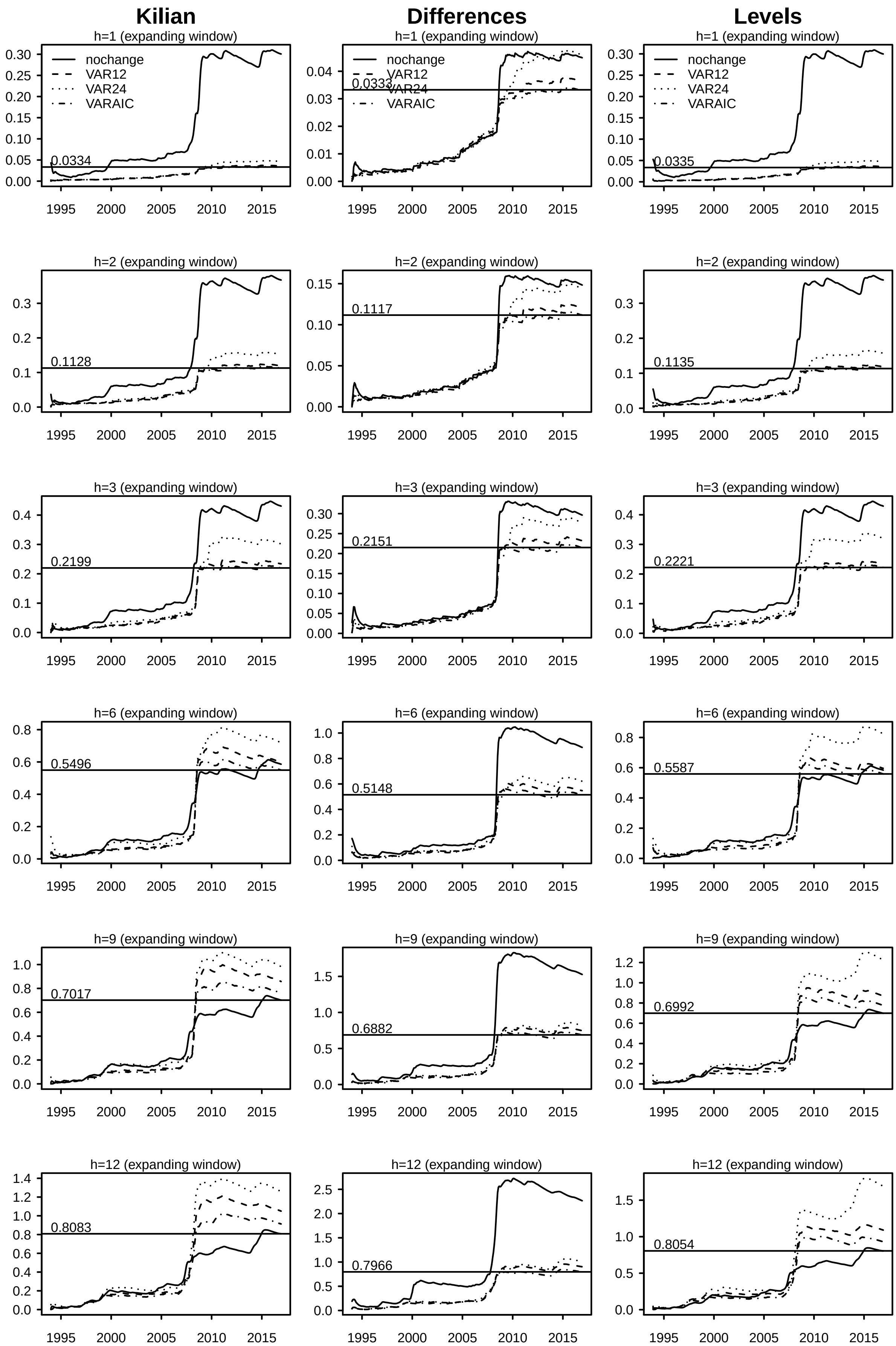
What we observe at first is the general tendency of a steady increase of the MSE over time. Thus the accuracy of the oil price forecasts deteriorates systematically since the 1990s. This might be explained by the several changes affecting global oil markets in the late 1980s (see Hamilton (2009)). The collapse of OPEC had lasting implications.

[4] Some corresponding results with a rolling window design of the forecast experiment (in fact a rolling window of 240 months) are collected in the appendix.

[5] Depicted is $MSE_t = t^{-1}\sum_{s=1}^{t}(y_s - \hat{y}_{s,h})^2$ with the realization of the real oil price denoted by y_s (not in logs) and the h-step forecast $\hat{y}_{s,h}$ for the same period s, obtained by a particular method (indicated in the legend in the first row of the figure) and appropriately retransformed from of the variables included in the VAR.

[6] Using the Bayesian Information Criterion (BIC) leads to very similar lag length selection and very similar results.

Figure 4: Benchmark Selection (expanding window)



The powerful cartel from the 1970s never recovered from the oil price collapse in 1986 and permanently lost influence on global markets. The fall of the Soviet Union and the emergence newly independent oil producing countries was a further source of oil market disruptions. The second obvious characteristic is the impact of the financial and economic crisis with the consequence of a series of particularly bad forecasts, leading to a pronounced rise of the MSE lines.[7] After about 2010 forecast errors stabilize on a high level or appear to to improve by a small margin.

The central column of the figure clearly shows that MSE values obtained with a VAR in differences are generally smaller than those obtained with the Kilian VAR and the VAR in levels across all specifications. However, looking down the columns of the figure we observe that the forecast performance quickly deteriorates with increasing forecast horizon. For the VAR in differences the VARAIC is the best forecasting method, closely followed by VAR12 and VAR24. The left and right columns in the figure, pertaining to the Kilian VAR and the VAR in levels, respectively, roughly contain the same message. For the shorter forecast horizons ($h = 1, 2, 3$) the VARAIC performs better than the VARs with a lag order fixed at 12 or 24, while the no-change forecast performs worst. In contrast, for the longer forecast horizons ($h = 6, 9, 12$), the VAR12 and VAR24 perform poorly, while there is a close competition of the VARAIC and the naïve no-change forecasts with the no-change forecasts becoming slightly better at the longest forecast horizons.

Taking these results together we select the VARAIC as the overall best forecasting method and decide to use this method as the benchmark in the subsequent comparison with the sparse VAR approaches.[8] This allows for a direct comparison of the effects of the regularization (imposing sparsity) within the common framework of a VAR model. The main issue is the distinction of pruning entire coefficient matrices versus pruning single coefficients within these matrices.

4.2 Sparse VARs

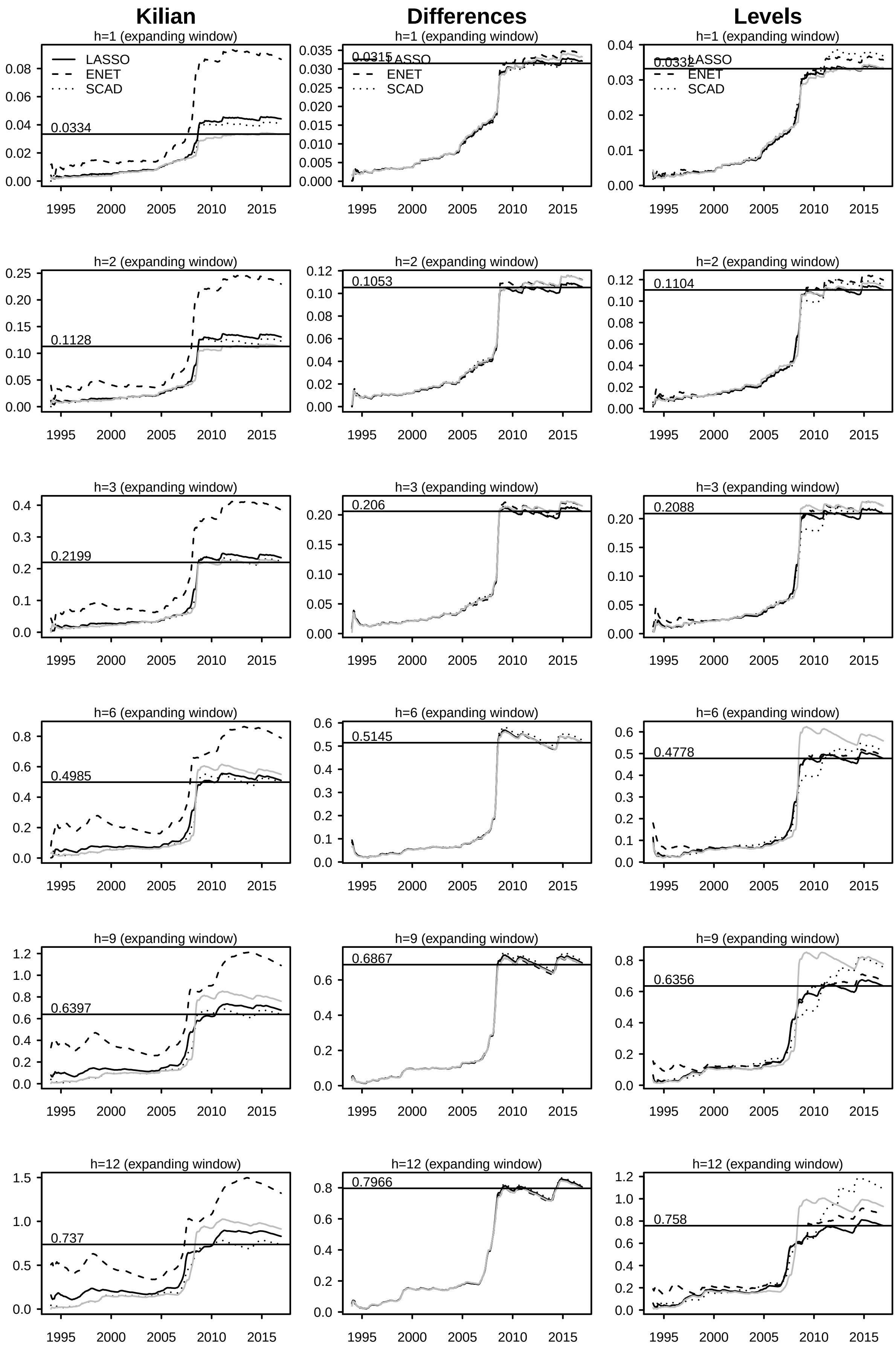
Figure 5 is an analogous depiction of the results for the sparse VAR models, i.e. the basic LASSO, ENET and SCAD, shown by the solid, dashed and dotted black lines, respectively. The recursive MSE of the benchmark VARAIC is shown as gray lines. We start with a VAR(12) to which the regularization is applied. Note that the regularization in equation (6) depends on the relative magnitudes of the parameters which in turn depends on the scaling of the variables. Thus, all variables are normalized to have the same standard deviation, which is the standard deviation of the log real oil price.

As before, we find the same general increase of the forecast error measures, especially during the months of the financial crisis. Skimming through the forecast horizons in search for the best combination of variable transformation and estimation method we first of all observe that the forecasts from a VAR in differences remain slightly better than those obtained from the VAR with variables transformed according to Kilian and the VAR in levels at shorter forecast horizons but loses ground at longer forecast horizons. This holds irrespective of the particular form of regularization used for the VAR estimation.

[7] This large increase of the forecast errors is also visible in the related plots of Zhang et al. (2019, p. 105).

[8] We also estimated vector error correction models (VECM) with imposing cointegration relations determined by the Johansen (1988, 1991) methodology. The forecast errors do not point to an improvement of the predictive performance compared to the VARAIC.

Figure 5: Evaluation of the Sparse VARs (expanding window)



Comparing the Kilian VAR and the VAR in levels we find that the MSE at the end of the forecast period obtained with the best method (the number above the horizontal line) is smaller in the case of the VAR in levels for all forecast horizons. The particular estimation method which reaches the smallest MSE at the end differs, however. For the shorter forecast horizons ($h = 1, 2, 3$) and the Kilian VAR the VARAIC is best, closely followed by SCAD and LASSO, while ENET performs worst. In the case of the VAR in levels the ranking is different. Here, the LASSO and the ENET are the best methods and are close to each other. The VARAIC is also close for the shortest forecast horizons with a widening gap when the forecast horizon increases. The SCAD method performs worst for all forecast horizons.

For the longer forecast horizons ($h = 6, 9, 12$) the LASSO and ENET are best in the VAR in levels and overall. For the VAR in levels the VARAIC and SCAD get worse with increasing forecast horizon. Interestingly, VARAIC and SCAD perform best before the financial crisis and also perform better than the other methods when the forecast performance worsens during the period of the financial crisis. SCAD remains best until about 2010 as is visible by the dotted line being lower than the other lines. The ENET is the worst method before the financial crisis and gains much in performance afterwards. Again, there are differences when the Kilian VAR is considered. With this transformation, SCAD and LASSO are best at the end of the sample period and also perform quite well before, in particular since the financial crisis. While VARAIC and SCAD are best performing before the financial crisis, VARAIC loses much more performance during the financial crisis than SCAD does. The ENET generally performs worst in the Kilian VAR.

Considering all results together, we see that the forecast performance depends on the transformation of the variables considered in the VAR and there are also pronounced differences across all employed regularization approaches to induce sparsity. Most important, if a particular regularization method performs well with a particular transformation of the variables this does not imply that the same method also performs well with a different variable transformation. The comparison of ENET and SCAD shows this clearly. The basic LASSO appears to be a quite good allrounder which not always performs best but adapts well to different transformations of the variables and is never far behind the best performing method.

5 Extended Variable Sets

One of the main virtues of sparse regression methods is the property to deal with situations in which there are more variables than observations. This is enabled by the LARS algorithm (least angle regression, Efron et al. (2004)) which can cope with those situations (see Hastie et al. (2009, ch. 18) for an exposition). When we extend the variable set consisting of the three core variables considered so far by further variables the number of parameters grows with the square of the number of variables in the VAR for a constant lag length. Thus, it is of particular interest to investigate whether the sparse VAR methods are able to exploit the predictive power of further variables in extended variable sets.

In this section the VAR model with the three core variables is extended by different variable sets containing industrial production indices of the G7 countries[9], exchange rates

[9]Canada, France, Germany, Italy, Japan, United Kingdom and United States.

to the US dollar and variables related to different investment opportunities. Further more, we also apply impulse indicator saturation (IIS) to eliminate the potentially adverse effects of single observations on the forecast performance.

The panels in subsequent figures are arranged analogous to the previous section. Now the black lines show the cumulative MSE values with the extended variable sets (solid for the LASSO, dashed for ENET and dotted for SCAD). The gray lines represent the corresponding results only including the three core variables as discussed above (see figure 5) for the purpose of a direct comparison of the effects of the enlarged variable sets.

5.1 Production Indices

The theoretical reasoning behind these additional variables is obvious in the case of the industrial production (or changes thereof) which is a major driver of the oil price. Admittedly, large newly industrializing countries like China or India are not included for reasons of data availability. Although this omission is not critical at the start of the sample period it may become increasingly crucial nearing the end of the sample period. As far as the industrial production in these countries is linked to the industrial production of the G7 countries this omission can be accommodated by the VAR coefficients. The data for industrial production are retrieved from the FRED database.[10] Under our three transformations of the variables we take log differences of the production indices in the case of the Kilian VAR and the VAR in differences, whereas we use log levels in the case of the VAR in levels.

Adding the production indices of the G7 countries leads to the results shown in figure 6 under an expanding window design. Considering first the Kilian VAR with the growth rates (computed as log differences) of the production indices added we observe that the smallest final MSE values are reached by the SCAD method for all forecast horizons considered. The results are almost indistinguishable from the previous results without including the production indices (in the figure the gray dots for the SCAD results are almost completely plotted over the black dots from the extended model). It seems that the additional variables are completely pruned out by SCAD regularization. The other regularization methods, i.e. LASSO and ENET, are associated with larger final MSE values when the production indices are included.

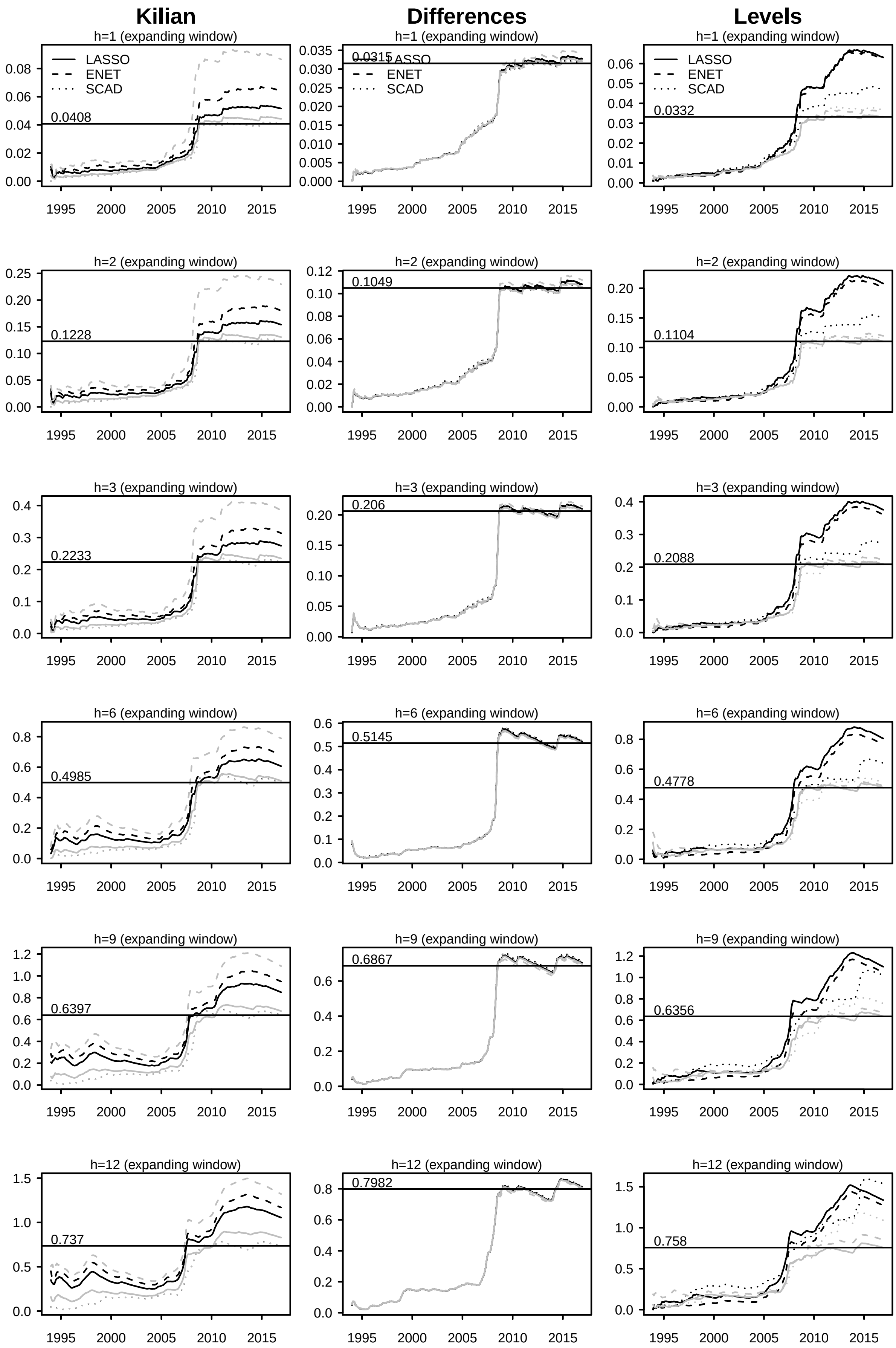
In the case of the VAR in differences we observe no smaller MSEs across all forecast horizons compared to the results of the previous section. We also find substantially larger MSE values across all horizons when the VAR in levels is extended by the log levels of the production indices. Here, the increase in MSE is so large that the ranking of the Kilian VAR and the VAR in levels reverses. Specifically, the Kilian VAR is now better than the VAR in levels in terms of final MSE (black lines) but is not better than the VAR in levels without the extension by the production indices (gray lines).

5.2 Exchange Rates

In parallel with the previous subsection we now extend the variable set by the exchange rates of the G7 countries (excluding the US) to the US dollar. Since the international

[10]The respective codes are CANPROINDMISMEI, FRAPROINDMISMEI, DEUPROINDMISMEI, ITAPROINDMISMEI, JPNPROINDMISMEI, GBRPROINDMISMEI and INDPRO.

Figure 6: Sparse VARs Augmented with Industrial Production (expanding window)



oil trade is conducted in US$, it makes sense to extend the VAR by the exchange rates of the G7 countries (excluding the US) to the US dollar. This is especially true when we consider that most G7 countries heavily rely on imported oil traded in US$ to satisfy domestic demand. The theoretical justification for adding the exchange rates is based on models such as Krugman (1983a,b). For empirical work on the relationship between crude oil prices and real exchange rates we refer to Zhou (1995), Amano and van Norden (1995) and Bénassy-Quéré et al. (2007). The series on the exchange rates are also taken from the FRED database.[11] We apply the same transformations as in the case of the production indices in the previous subsection.

The results are reported in figure 7. The extension of the Kilian VAR by the log differences of the exchange rates gives rise to similar results as the extension by industrial production in the previous subsection. In the case of the the VAR in differences we observe no improvement and the VAR in differences remains the worst performing model for the longer forecast horizons with larger forecast errors than the Kilian VAR and the VAR in levels. Finally, in the VAR in levels, extended by the log exchange rates, there is a substantial deterioration of forecast performance of all three sparse estimation methods across all forecast horizons. Here now the performance of SCAD also deteriorates.

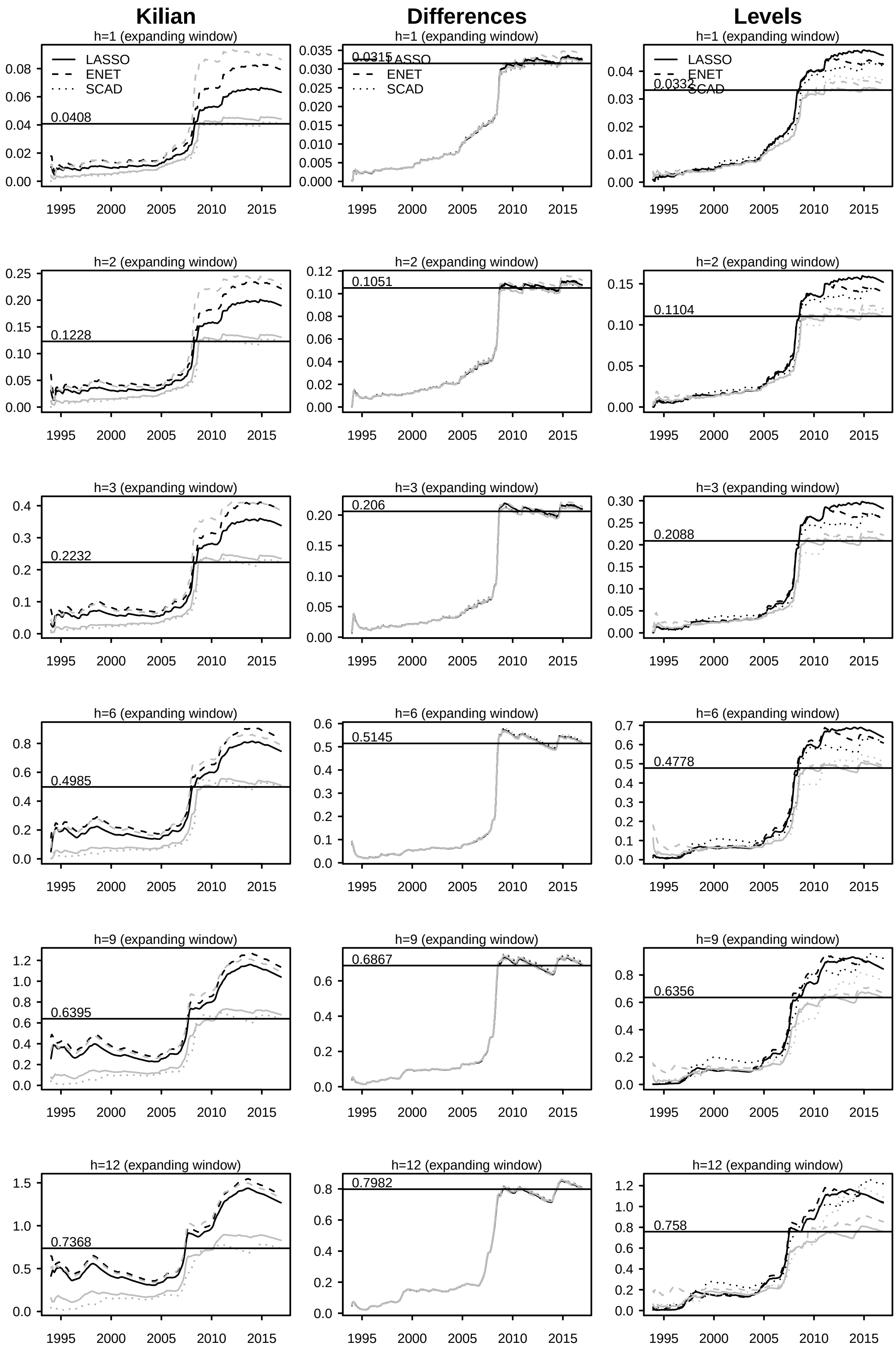
Taken together, we find no improvement by augmenting the VARs with the production indices or exchange rates and applying the sparse estimation methods to eliminate unimportant variables and lags. There are two aspects leading to this outcome. The first possibility is that the sparse VAR estimators are not able to filter out the relevant variables and lags. Given that, the sparse VAR methods apparently fail to set the parameters to zero which actually are equal to zero leading to more noisy parameter estimates and forecasts, finally resulting in larger MSE values. The second possibility is that the variables used for extending the model are largely irrelevant for forecasting the world oil price or contain information which is already comprised in the three core variables. In the case of the production indices it seems quite plausible that they represent information about economic activity in the G7 countries which is also contained in the global real activity index. This is, however, not born out by the correlations of the production indices with the real activity index (max. correlation ≈ 0.12 with real activity and ≈ 0.38 with changes of real activity). Here it is important to recall that the real activity index is a global measure based on international dry cargo shipping rates and therefore also comprises the activity of other large emerging economies like China and India.

5.3 Investment Opportunities

As a third extended variable set we consider the prices of different investment opportunities as possible candidate variables. There is a literature on the stock market effects of oil price shocks (see Kilian and Park (2009) among many others). Here, the identifying restrictions imposed postulate the instantaneous response of a stock market index (actually the real log returns of the CRSP value-weighted market portfolio), but not the other way round. In a VAR there may be, however, also a response of the oil price to the stock

[11] The respective codes are EXCAUS, CCUSMA02FRM618N, CCUSMA02DEM618N, CCUSMA02ITM618N, CCUSMA02GBM618N and EXJPUS. For France, Germany and Italy we point out, that exchange rate is expressed in Euro to US$. Before the introduction of the common currency in January first, 1999 the series are constructed by using the official national fixed exchange rates to the Euro.

Figure 7: Sparse VARs Augmented with Exchange Rates (expanding window)



market reaction in the next period. Thus, the information comprised in the returns of different investment opportunities may be suitable to improve the oil price forecasts. Those forward-looking variables are also considered in business cycle research (see e.g. Stock and Watson (2003)). The usefulness of financial market data in forecasting oil prices is also subject of recent research (e.g. Degiannakis and Filis (2018), Zhang et al. (2019)), This is also acknowledged by Zhang et al. (2019, p. 108) stating that "an increasing number of financial institutions view crude oil as a new class of financial asset and start to invest in the crude oil market to diversify their portfolios. This leads to a strong correlation between the stock and oil markets."

To assess this issue we include the index values or the returns of CSRP market portfolio[12], the real gold price[13], a comprehensive bond price index[14] in the sparse VAR models. These time series are transformed by logs and are differenced in the cases of the Kilian VAR and the VAR in differences. In addition, the 3 month and 10 year treasury rates[15] are also included without transformation.

The relevance of the additional financial variables is mainly motivated by the arbitrage condition linking the crude oil spot price to crude oil futures prices (see Fatouh et al. (2012)). As Hamilton and Wu (2014) point out, the futures market started to expand very quickly in the early 2000s, primarily due to crude oil futures viewed as an instrument for portfolio diversification. Thus, we include variables that have frequently been used to determine returns in futures markets (see e.g. Bessembinder (1992), De Roon et al. (2000) and Hong and Yogo (2012)).

The results are shown in figure 8. As before in the case of the introduction of the production indices and the exchange rates we find no improvement of the forecast performance, in particular since the financial crisis and the great recession. In the Kilian VAR we observe a deterioration of the forecasts based on LASSO and ENET estimates across all forecast horizons. Only SCAD achieves the same performance as the basic model with the three core variables (again shown as gray lines for reference), most likely caused by the total elimination of the additional variables. Considering the VAR in differences we find the forecast errors achieved with the additional variables to be very similar to those without augmentation. The VAR in levels leads to worse forecasts for all augmented models across all horizons. This holds in particular for the SCAD estimated models at larger forecast horizons.

5.4 Impulse Indicator Saturation

The final attempt to improve the accuracy of the oil price forecasts is the introduction of impulse indicator saturation (IIS) in the estimation procedure of the VARs. This device, introduced by Santos et al. (2008), uses a complete set of dummy variables (one for each observation) to prune out single observations from the whole data series

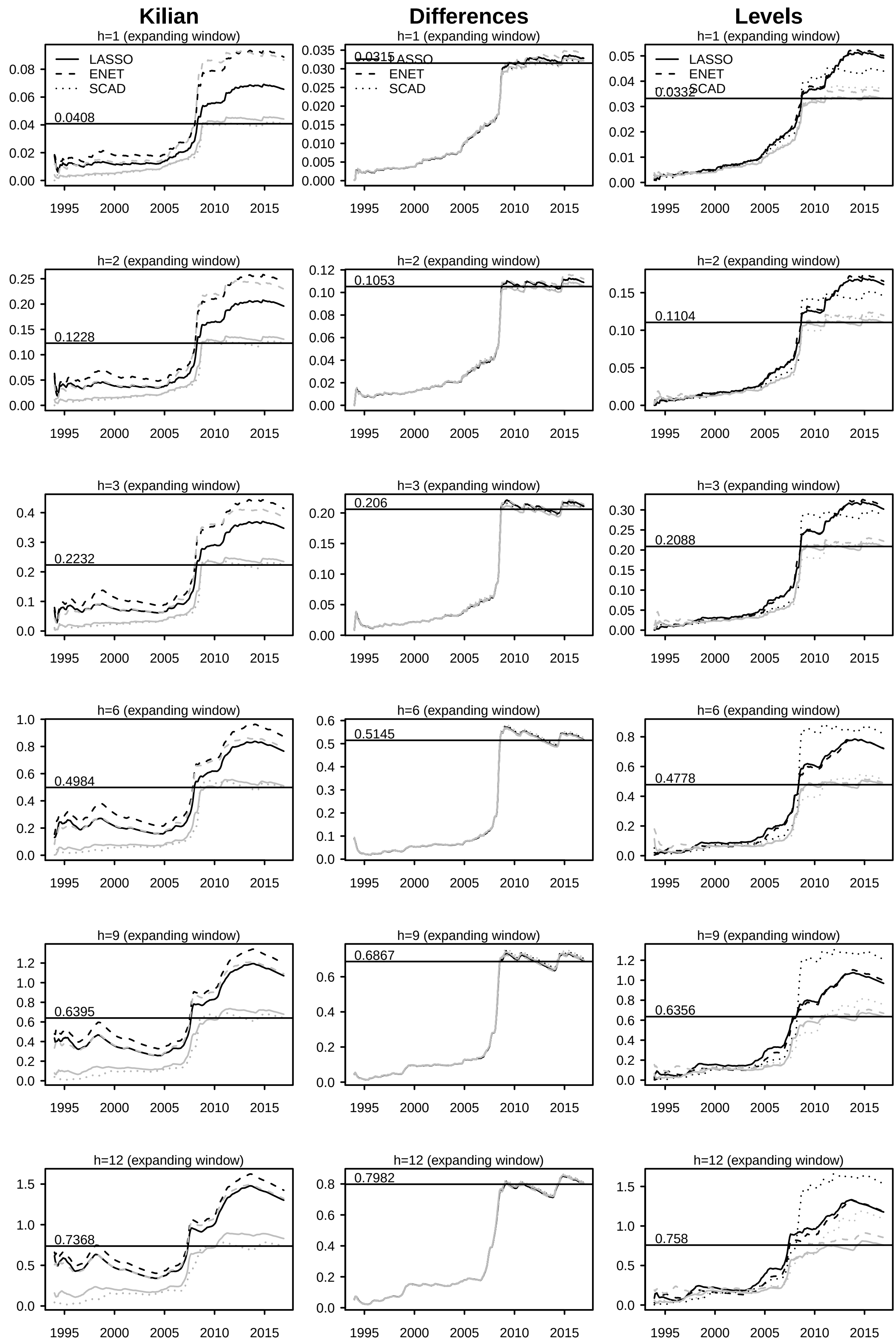
[12]This is a broad value-weighted index as the market portfolio, formed on the universe of all CRSP firms incorporated in the US. Data are from the data archive of Kenneth French.

[13]Price of one fine ounce in US dollar, daily fixed at the London Bullion Market, averaged over the respective month and transformed to a real price by the US CPI. Data are from the time series database of the Deutsche Bundesbank.

[14]This is the BofA Merrill Lynch US Corp Master Total Return Index Value, transformed to a real price by the US CPI.

[15]Both series are retrieved from the FRED database with the corresponding codes GS3M and GS10.

Figure 8: Sparse VARs Augmented with Investment Opportunities (expanding window)



which may represent outliers or result from structural breaks. Since the complete set of dummy variables cannot be introduced into the VAR at once, the first step of the procedure includes the dummy variables for the first half of the sample period and tests which of them are significant on a 5 percent level. In the second step, only the dummy variables for the second half of the sample period are included and tested in the same way. Finally, all dummy variables which have been found significant in the first and the second step are introduced simultaneously and again individually tested for their significance. The subset of those dummy variables which remain significant in the final step are then kept in the VAR for the estimation. This actually amounts to exclude the associated observations. This procedure is conducted anew in each forecast step of our forecast evaluation procedure.

Figure 9 shows the results when all t-tests in the IIS procedure are performed on a 5 percent level of significance. Compared are the VAR with 12 lags estimated by OLS (VAR12) or subjected to the selection of variables by the Lasso operator (LASSO), represented by gray lines, with their variants estimated after performing the IIS (denoted IISVAR12 and IISLASSO), represented by black lines. In the case of the VAR(12) the application of IIS does not lead to an improvement of the forecasts and even leads to deteriorations at the longer forecast horizons. This holds likewise for the Kilian transformation and for the VAR in levels, whereas the curves in the case of the VAR in differences are very close.

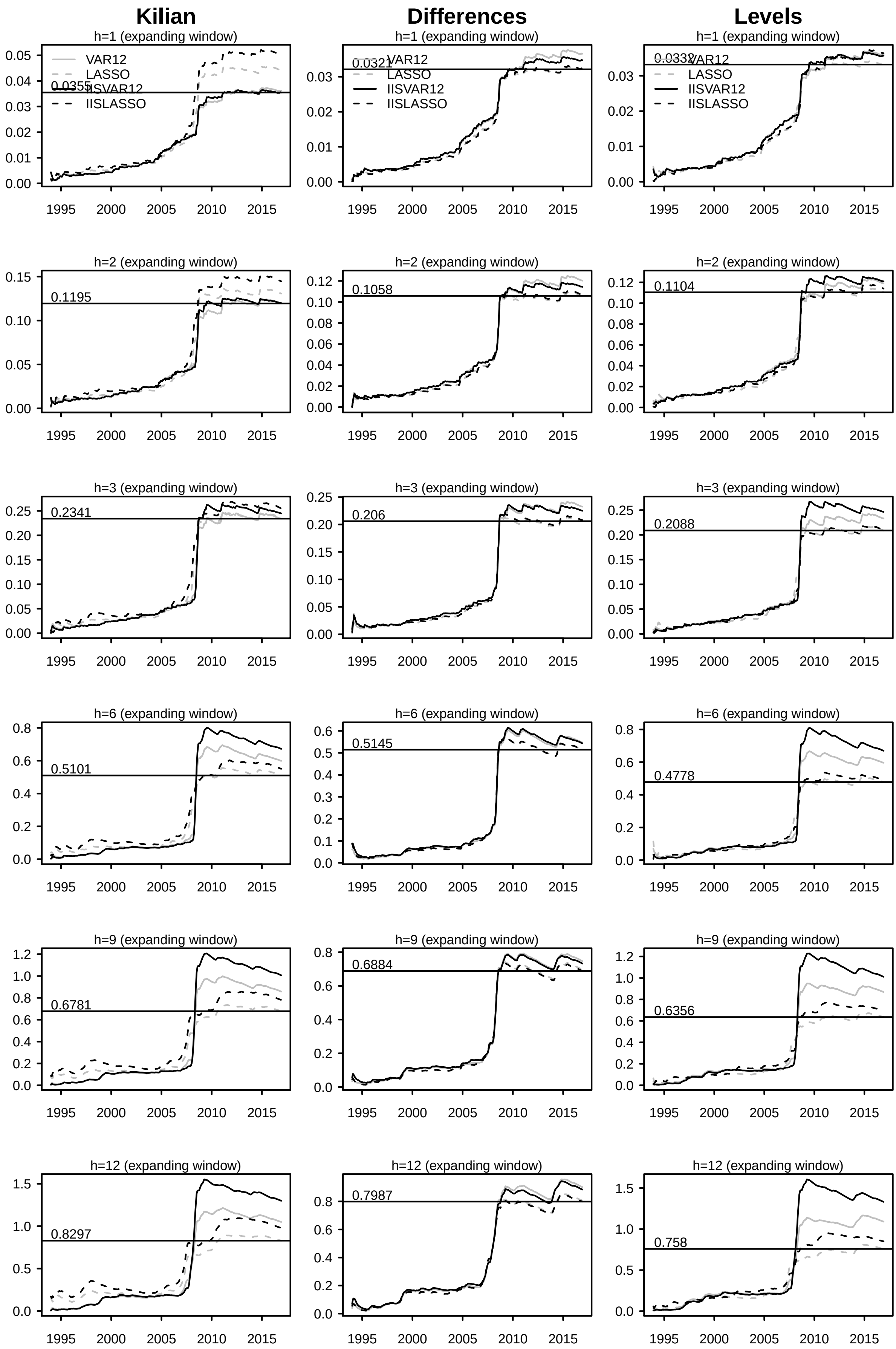
Comparing the LASSO-based forecasts with and without the IIS in advance we see that those without IIS nearly always have an edge over those with IIS. Again, the difference becomes larger with increasing forecast horizon for the Kilian transformation and the VAR in levels and is negligible for the VAR in differences. Nevertheless, both LASSO-based forecasts are not far away from the ordinary VAR forecasts before the financial crisis but become much better thereafter. Repeating the analysis with a 1 percent level of significance leads to results (not shown) that are less favorable for the IIS in this application. Finally, turning to a rolling window instead of an expanding window for the estimation and IIS selection leads to considerably larger forecast errors when using IIS (see the appendix).

In sum, we can draw the conclusion of the analysis with the extended information sets in this section that the VAR in levels with just the three core variables estimated by sparse VAR methods (as discussed in section 4.2 above) remains the best overall forecast method.

6 Conclusion

In the above analysis we have conducted a forecast evaluation exercise for the world real price of crude oil. Our point of departure was the three-variable VAR model of Kilian (2009) which has already been subjected to a forecast evaluation by Alquist et al. (2013). The value added of our analysis is the application of estimation methods based on regularization to achieve sparsity in the parameter matrices of the VARs. Whereas classical information criteria for lag-order selection such as AIC or BIC restrict entire parameter matrices for lag orders higher than the selected to zero, the regularization methods have the property to restrict specific parameters within the parameter matrices to zero while others retain their values different from zero. By that, the detrimental effect of parameter estimates which are truly zero but are estimated with small magnitudes and

Figure 9: Evaluation with IIS (expanding window)



are likely to be insignificant on forecast performance is reduced. This holds the promise of reaching a better forecast performance.

The main results of this forecast evaluation exercise can be summarized in three main lessons. The first lesson is that the selection of the benchmark VAR reveals that "long" VARs including many lags of the variables (e.g. 12 or 24) has a justification for impulse response analysis, but is detrimental to forecast performance. As emphasized in Kilian and Lütkepohl (2017, pp. 63ff.) covering a cycle of a year with monthly data (or a multiple thereof) is important for impulse response analysis. However, we have seen that these "long" VARs are clearly dominated by more parsimonious VARs with respect to forecasting performance. The second lesson is that the application of regularization of the VARs for improving forecast performance depends on the choice of variable transformations. We find that regularization improves forecasts especially for the longer forecast horizons up to 12 months for the VAR with variables transformed according to Kilian (2009) and the VAR in levels. The forecasts for the VAR in differences are also improved by applying the sparse estimators, but here only for the shorter forecast horizons. The third lesson is that extending the variable set and then applying the sparse VAR estimators does not generally lead to further reductions of the forecast errors. Thus, the general property of the LASSO and related estimators as devices for the selection of suitable sets of predictors out of a large set of candidate variables as supposed for simple linear regression is not born out in the case of VARs. Instead, we find that the forecast performance of the augmented VARs worsens or the additional variable-lag combinations are totally pruned out by the sparse estimators.

This outcome stands in contrast to other related macroeconomic forecast evaluation exercises (not focusing on the oil price) such as Nicholson et all. (2017). Therefore, it appears that more experience with the regularization methods for estimating sparse VARs in different situations is required. This paper makes a contribution to this endeavor. Along these lines an investigation of the suitability of other regularization approaches for VARs such as the variants recently proposed by Nicholson et al. (2016, 2017) would be valuable. The current implementation in the R-package "BigVAR" appears rather slow in terms of computation time, however. This may not be a problem for computing a single or a small number of forecasts, but it becomes prohibitive for a forecast evaluation exercise where hundreds or thousands of forecasts need to be computed as we have done in this paper. Hence, we have not applied these methods in the present paper, but look forward to do so in the future when faster computers and/or software are available. Finally, it would be interesting to have a closer look into the estimation results to investigate which variable-lag combinations are eliminated by the regularization and to see whether is pattern is stable over time or is subject to systematic changes. This is beyond scope of the present paper, but is an interesting opportunity for future research.

Appendix: Results Under a Rolling Window Design

The figures A1, A2 and A3 in this appendix show the results with a rolling window design of the forecast experiment. Used is a rolling window of 240 months for the estimation sample.

Summarizing these results we can conclude that the overall pattern of results is quite similar when using the rolling window instead of the expanding window. However, we generally find larger forecast errors and larger final MSE values at the end of the evaluation period. Using the sparse VAR methods results in smaller improvements for the shorter forecast horizons and no visible improvements and even deteriorations at the longer horizons. In particular, for the VAR in levels we find a tremendous deterioration of LASSO and ENET at longer horizons in the aftermath of the financial crisis which does not occur when using the expanding window. Invoking the IIS procedure does only occasionally lead to improvements of the forecast performance and generally causes the forecast errors to be larger.

Overall and irrespective of the variable transformations we find a better forecast performance at the end of the sample period with the expanding window instead of the rolling window. This outcome may be a cause of the larger sample size available when using the expanding window. A major contribution can attributed to the financial crisis which is comprised in each of the rolling window estimation samples until the end of the evaluation period with a larger weight than in the expanding windows.

Figure A1: Benchmark Selection (rolling window)

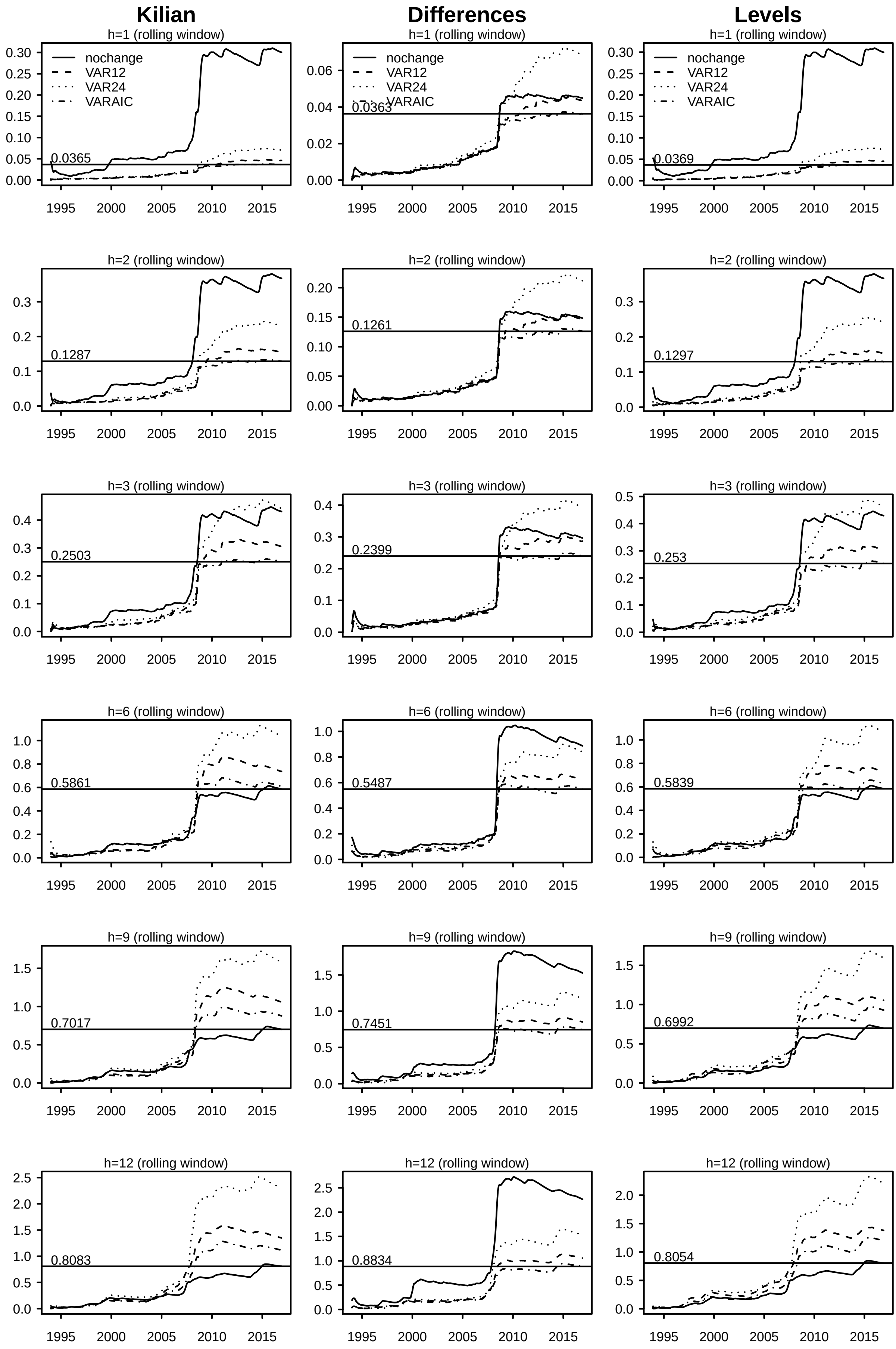


Figure A2: Evaluation of the Sparse VARs (rolling window)

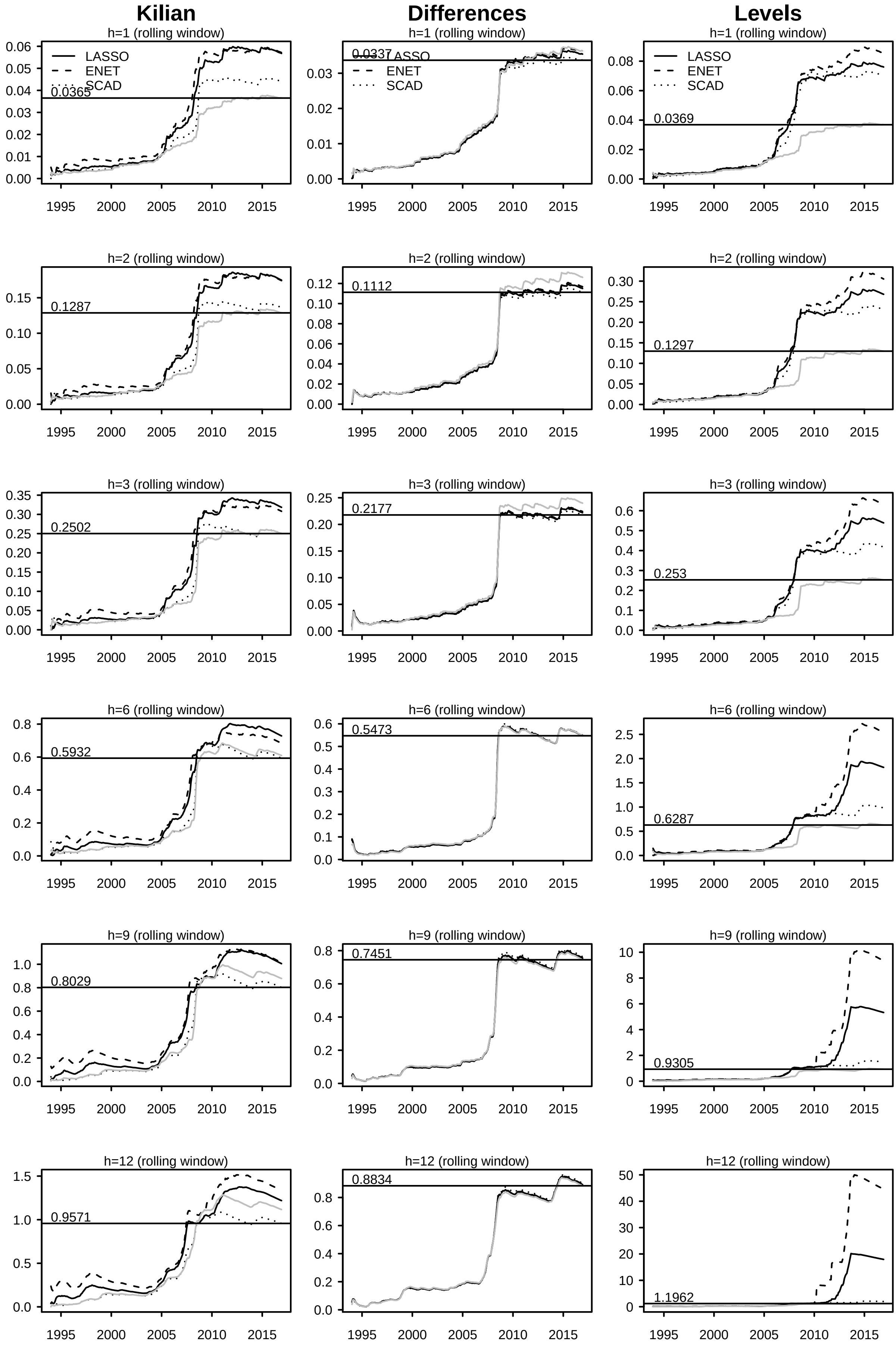
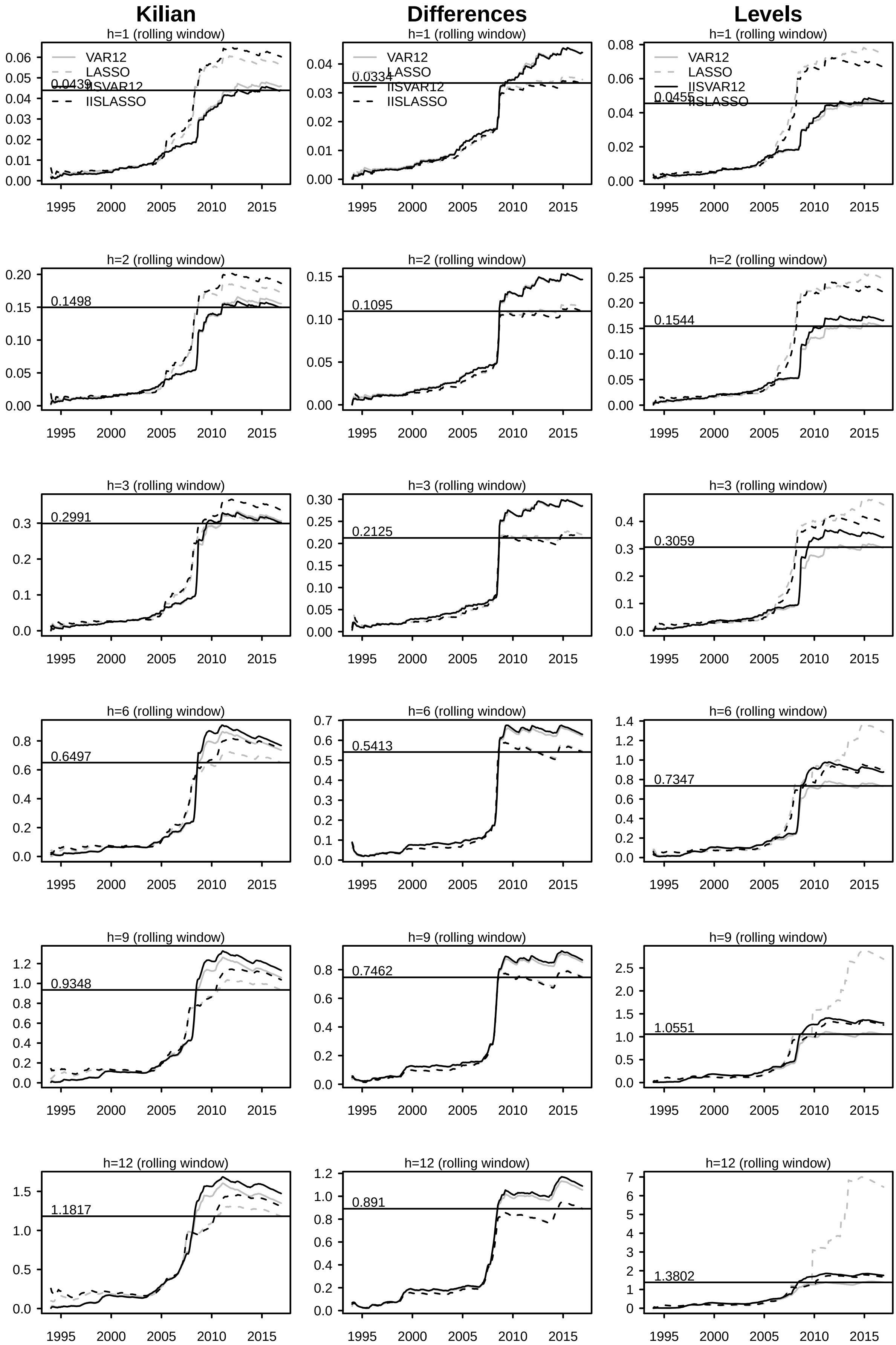


Figure A3: Evaluation with IIS (rolling window)



References

[1] Alquist, R., Kilian, L. Vigfusson, R.J. (2013), Forecasting the Price of Oil, in: G. Elliott, A. Timmermann (eds.), Handbook of Economic Forecasting, vol. 2A, Amsterdam: Elsevier, ch. 8, pp. 427–507.

[2] Amano, R. and van Norden, S. (1995), Exchange rates and Oil Prices, Bank of Canada, Working Paper.

[3] Baumeister, C., Kilian, L. (2014), What Central Bankers Need to Know about Forecasting Oil Prices, International Economic Review 55, 869–889.

[4] Baumeister, C., Kilian, L. (2015), Forecasting the Real Price of Oil in a Changing World: A Forecast Combination Approach, Journal of Business and Economic Statistics 33, 338–351.

[5] Baumeister, C., Kilian, L. (2016), Forty Years of Oil Price Fluctuations: Why the Price of Oil May Still Surprise Us, Journal of Economic Perspectives 30, 139–60.

[6] Baumeister, C., Kilian, L., Lee, T.K. (2014), Are There Gains from Pooling Real-Time Oil Price Forecasts?, Energy Economics 46, S33–S43.

[7] Bénassy-Quéré, A., Mignon, V. and Penot, A. (2007), China and the relationship between the oil price and the dollar, Energy Policy 35, 5795–5805.

[8] Bessembinder, H. (1992), Systematic Risk, Hedging Pressure, and Risk Premiums in Futures Markets, Review of Financial Studies 4, 637–667.

[9] Cheng, F., Li, T., Wei, Y., Fan, T. (2019), The VEC-NAR Model for Short-Term Forecasting of Oil Prices, Energy Economics, vol. 78, pp. 656–667.

[10] Christiano, L.J., Eichenbaum, M., Evans, C.L. (2005), Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy, Journal of Political Economy 113, 1–45.

[11] Drachal, K. (2016), Forecasting Spot Oil Price in a Dynamic Model Averaging Framework: Have the Determinants Changed over Time?, Energy Economics 60, 35–46.

[12] Degiannakis, S., Filis, G. (2018), Forecasting Oil Prices: High-frequency Financial Data are Indeed Useful, Energy Economics 76, 388–402.

[13] De Roon, F.A., Nijman, T. E., Veld, C. (2000), Hedging Pressure Effects in Futures Markets The Journal of Finance, Blackwell Publishers 55, 1437–1456.

[14] Efron, B., Hastie, T., Johnstone, I., Tibshirani, R. (2004), Least Angle Regression (with discussion), Annals of Statistics 32, 407–499.

[15] Elliott, G., Rothenberg, T.J., Stock, J.H. (1996), Efficient Tests for an Autoregressive Unit Root, Econometrica 64, 813–836.

[16] Engle, R.F., Granger, C.W.J. (1987), Co-Integration and Error Correction: Representation, Estimation and Testing, Econometrica, 55, 251–276.

[17] Fan, J., Li, R. (2001), Variable Selection via Nonconcave Penalized Likelihood and its Oracle Properties, Journal of the American Statistical Association 96, 1348–1360.

[18] Fattouh, B., Kilian, L., Mahadeva, L. (2013), The Role of Speculation in Oil Markets: What Have We Learned So Far?, Energy Journal 34, 7–33.

[19] Funk, C. (2018), Forecasting the Real Price of Oil: Time-Variation and Forecast Combination, Energy Economics 76, 288–302.

[20] Garratt, A., Vahey, S.P., Zhang, Y. (2019), Real-Time Forecast Combinations for the Oil Price, Journal of Applied Econometrics 34, 456–462.

[21] Hamilton, J.D. (2018), Measuring Global Economic Activity, https://econweb.ucsd.edu/~jhamilto/REA.pdf.

[22] Hamilton, J.D. (2009), Understanding Crude Oil Prices, Energy Journal, 30. 179–206.

[23] Hamilton, J.D., Wu, C. (2014), Risk Premia in Crude Oil Futures Prices, Journal of International Money and Finance 42, 9–37.

[24] Hastie, T., Tibshirani, R., Friedman, J. (2009), The Elements of Statistical Learning: Data Mining, Inference, and Prediction, second edition, New York: Springer.

[25] Hoerl, A.E., Kennard, R. (1970), Ridge Regression: Biased Estimation for Nonorthogonal Problems, Technometrics 12, 55–67.

[26] Hoffman, D.L., Rasche, R.H. (1996), Assessing Forecast Performance in a Cointegrated System, Journal of Applied Econometrics 11, 495–517.

[27] Hong, H., Yogo, M. (2012), What Does Futures Market Interest Tell Us About the Macroeconomy and Asset Prices? Journal of Financial Economics, 105, 473–490.

[28] Johansen, S. (1988), Statistical Analysis of Cointegration Vectors, Journal of Economic Dynamics and Control 12, 231–254.

[29] Johansen, S. (1991), Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models, Econometrica 59, 1551–1580.

[30] Kilian, L. (2019), Measuring Global Real Economic Activity: Do Recent Critiques Hold Up to Scrutiny?, https://sites.google.com/site/lkilian2019/home.

[31] Kilian, L. (2009), Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market, American Economic Review 99, 1053–1069.

[32] Kilian, L. (2008), The Economic Effects of Energy Price Shocks, Journal of Economic Literature 46, 871–909.

[33] Kilian, L., Hicks, B. (2013), Did Unexpectedly Strong Economic Growth Cause the Oil Price Shock of 2003–2008?, Journal of Forecasting 35, 385–394 .

[34] Kilian, L., Lütkepohl, H. (2017), Structural Vector Autoregressive Analysis, Cambridge (Mass.): Cambridge University Press.

[35] Kilian, L., Park, C. (2009), The Impact of Oil Price Shocks on the U.S. Stock Market, International Economic Review 50, 1267–1287.

[36] Krugman, P. (1983a), Oil and the Dollar, in: J. Bhandari, B. Putnam (eds), Economic Interdependence and Flexible Exchange Rates, Cambridge (Mass.): MIT Press, pp. 179–190.

[37] Krugman, P. (1983b), Oil Shocks and Exchange Rate Dynamics, in: J.A. Frankel (ed), Exchange Rates and International Macroeconomics, Chicago and London: University of Chicago Press, pp. 259–283.

[38] Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y. (1992), Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root, Journal of Econometrics 54, 159–178.

[39] Lütkepohl, H. (2005), New Introduction to Multiple Time Series Analysis, Springer: Berlin.

[40] Murphy, K.P. (2012), Machine Learning: A Probabilistic Perspective, Cambridge (Mass.): MIT Press.

[41] Naser, H. (2016), Estimating and Forecasting the Real Prices of Crude Oil: A Data Rich Model Using a Dynamic Model Averaging (DMA) Approach, Energy Economics 56, 75–87.

[42] Nicholson, W.B., Bien, J., Matteson, D.S. (2016), High Dimensional Forecasting via Interpretable Vector Autoregression, arXiv:1412.5250v2.

[43] Nicholson, W.B., Matteson, D.S., Bien, J. (2017), VARX-L: Structured Regularization for Large Vector Autoregressions with Exogenous Variables, International Journal of Forecasting 33, 627–651.

[44] Santos, C., Hendry, D.F., Johansen, S. (2008), Automatic Selection of Indicators in a Fully Saturated Regression, Computational Statistics 23, 317–335.

[45] Sims, C.A. (1980), Macroeconomics and Reality, Econometrica 48, 1–48.

[46] Sims, C.A., Stock, J.H., Watson, M.W. (1990), Inference in Linear Time Series Models With Some Unit Roots, Econometrica 58, 113–144.

[47] Stock, J.H., Watson, M.W. (2003), Forecasting Output and Inflation: The Role of Asset Prices, Journal of Economic Literature 41, 788–829.

[48] Tibshirani, R. (1996), Regression Shrinkage and Selection via the Lasso, Journal of the Royal Statistical Society Series B, 58, 267–288.

[49] Wang, Y., Liu, L., Wu, C. (2017), Forecasting the Real Prices of Crude Oil Using Forecast Combinations over Time-Varying Parameter Models, Energy Economics 66, 337–348.

[50] Zhang, C.-H. (2010), Nearly Unbiased Variable Selection Under Minimax Concave Penalty, Annals of Statistics 38, 894–942.

[51] Zhou, S. (1995), The Response of Real Exchange Rates to Various Economic Shocks, Southern Economic Journal 61, 936–954.

[52] Zou, H., Hastie, T. (2005), Regularization and Variable Selection via the Elastic Net, Journal of the Royal Statistical Society B, 67, 301–320.

Arbeitspapiere der Volkswirtschaftlichen Fachgebiete der TU Darmstadt

ISSN: 1438-2733