On-line Appendix to Nonlinear Fore(Back)casting and Innovation Filtering for Causal-Noncausal (S)VAR Models

C. Gourieroux*, and J. Jasiak†

This version: May 21, 2023

*University of Toronto, Toulouse School of Economics and CREST, e-mail: *Christian.Gourieroux @ensae.fr*

†York University, e-mail: *jasiakj@yorku.ca*

On-Line APPENDIX B

Identification Conditions

B.1 Independent Component Analysis (ICA)

Let us consider the independent component model:

$$Y = D\epsilon, \tag{b.1}$$

where the observed vector Y is of dimension n and the components $\epsilon_1, ..., \epsilon_n$ are independent.

Proposition B.1. [Eriksson, Koivunen (2004), Th. 3, and Comon (1994), Th 11]

Under the following conditions:

i) D is invertible,

ii) the components $\epsilon_1, ..., \epsilon_n$ are independent and at most one of them has a Gaussian distribution,

the matrix D is identifiable up to the post multiplication by ΔQ, where Q is a permutation matrix and Δ a diagonal matrix with non-zero diagonal elements.

The matrix D is identifiable up to a permutation of indexes and to signed scaling $\epsilon_i \to \pm\sigma_i\epsilon_i$,with $\sigma_i > 0,\ i = 1, ..., n$. The only local identification issue is the positive scaling, which can be solved by introducing identifying restrictions.

Proposition B.2. [Hyvarinen et al. (2001)]

Under the assumptions of Proposition B.1. the local identification issue is solved if D is an orthogonal matrix: $D'D = Id$.

B.2 Two-Sided Multivariate Moving Averages

Proposition B.1 has been extended by Chan, Ho (2004), Chan, Ho, Tong (2006) to two-sided moving averages. We give a version of their result for structural mixed models:

$$Y_t = \Phi_1 Y_{t-1} + \cdots + \Phi_p Y_{t-p} + Du_t.$$

Proposition B.3

Let us assume that:

i) The roots of $\det(Id - \Phi_1 z - \cdots - \Phi_p z^p) = 0$ are not on the unit circle,

ii) Matrix D is invertible,

iii) (u_t) is i.i.d. with independent components,

iv) Each component admits a finite even moment of order k larger than 3, and at least one non-zero cumulant of order larger than 3.

Then i) $\Phi_1, ..., \Phi_p$ are identifiable, ii) D is identifiable up to the identification issues given in Proposition B.1.

This result corresponds to Condition 4 in Chan, Ho, Tong (2006). Assumption iv) implies that all distributions of the components are non-Gaussian.

The conditions of Proposition B.2. are sufficient for identification. Other sufficient conditions based on the cross-moments of 3rd and 4th order have been considered in the literature to weaken the assumption of cross-sectional independence [see Velasco (2022)].

References to On-Line Appendix B

Axler, S., Bourdon, P. and W. Ramey (2001): "Harmonic Function Theory", 2nd Ed., Graduate Texts in Mathematics, Springer.

Chan, K. and L. Ho (2004): "On the Unique Representation of Non-Gaussian Multivariate Linear Processes", Technical Report 341, University of Iowa.

Chan, K., Ho, L. and H. Tong (2006): "A Note on Time Irreversibility of Multivariate Linear Processes", Biometrika, 93, 221-227.

Comon, P. (1994): " Independent Component Analysis: A New Concept", Signal Processing, 36, 287-314

Eriksson, J. and V. Koivunen (2004): "Identifiability, Separability and Uniqueness of Linear ICA Models", IEEE Signal Processing Letters, 11, 601-604.

Hyvarinen, A., Karhunen, J. and E. Oja (2001): "Independent Component Analysis", Wiley.

On-Line Appendix C: Predictive Density Estimation

This Appendix describes the kernel-based estimation of the predictive density given in Proposition 1 from the following time series:

$$\hat{\epsilon}_t = Y_t - \hat{\Phi}_1 Y_{t-1} - \cdots - \hat{\Phi}_p Y_{t-p+1}, \quad t = 1, ..., T,$$

$$\hat{Z}_{2,t} = \hat{A}^2 \begin{pmatrix} Y_t \\ \tilde{Y}_{t-1} \end{pmatrix}, \quad t = 1, ..., T.$$

The above time series is used to approximate the density g of ϵ_t and density l_2 of $Z_{2,t}$ as follows:

$$\hat{g}_T(\epsilon) = \frac{1}{T} \frac{1}{h^m} \sum_{t=1}^{T} K_m \left(\frac{\epsilon - \hat{\epsilon}_t}{h} \right),$$

and

$$\hat{l}_{2,T}(z_2) = \frac{1}{T} \frac{1}{h^{n_2}} \sum_{t=1}^{T} K_{n_2} \left(\frac{z_2 - \hat{Z}_{2,t}}{h} \right),$$

where h_1, h_2 are the bandwidths and K_m, K_{n_2} are multivariate kernels of dimensions m and n_2, respectively. Then, the estimated predictive density is:

$$\hat{l}_T(y|\underline{Y}_T) = \frac{\hat{l}_{2,T} \left[\hat{A}^2 \begin{pmatrix} y \\ \tilde{Y}_T \end{pmatrix} \right]}{\hat{l}_{2,T} \left[\hat{A}^2 \begin{pmatrix} Y_T \\ \tilde{Y}_{T-1} \end{pmatrix} \right]} |\det \hat{J}_2| \, \hat{g}_T(y - \hat{\Phi}_1 Y_T - \cdots - \hat{\Phi}_p Y_{T-p+1}),$$

This formula is easily extended to bandwidths adjusted for each component, by replacing for example $\frac{1}{h^m} K_m \left(\frac{\epsilon - \hat{\epsilon}_t}{h} \right)$ by $\prod_{j=1}^{m} \frac{1}{h_j} K \left(\frac{\epsilon_j - \hat{\epsilon}_{j,t}}{h_j} \right)$, where K is a univariate kernel. Such an adjustment can account for different component variances.

Let us consider the example of a bivariate VAR(1) process with one noncausal component and a scalar noncausal eigenvalue λ_2 (see Section 6). The estimated coefficients of the inverse of A are denoted by:

$$\hat{A}^{-1} = \begin{pmatrix} \hat{a}^{11} & \hat{a}^{12} \\ \hat{a}^{21} & \hat{a}^{22} \end{pmatrix}.$$

The predictive density depends on unknown scalar parameters λ_1, λ_2 and functional parameters l_2, g that can be estimated. The marginal density $l_2(A^2 y)$ can be approximated by a kernel estimator:

$$\hat{l}_{2,T}(\hat{A}^2 y) = \frac{1}{T}\frac{1}{h_2}\sum_{t=1}^{T} K\left(\frac{\hat{a}^{21}(y_1 - y_{1,t}) + \hat{a}^{22}(y_2 - y_{2,t})}{h_2}\right),$$

while the density $l_2(A^2 Y_T)$, can be approximated by a kernel estimator:

$$\hat{l}_{2,T}(y_T) = \frac{1}{T}\frac{1}{h_2}\sum_{t=1}^{T} K\left(\frac{\hat{a}^{21}(y_{1,T} - y_{1,t}) + \hat{a}^{22}(y_{2,T} - y_{2,t})}{h_2}\right),$$

where h_2 is a bandwidth. The joint density $g(y - \Phi y_T)$ can be approximated by

$$\hat{g}_T(y - \hat{\Phi} y_T) = \frac{1}{T}\frac{1}{h_{11}h_{12}}\sum_{t=1}^{T} K\left(\frac{y_1 - \hat{\phi}_{1,1}y_{1,T} - \hat{\phi}_{1,2}y_{2,T} - \hat{\epsilon}_{1,t}}{h_{11}}\right) K\left(\frac{y_2 - \phi_{2,1}y_{1,T} - \phi_{2,2}y_{2,T} - \hat{\epsilon}_{2,t}}{h_{12}}\right).$$

where $\hat{\epsilon}_{1,t}$ and $\hat{\epsilon}_{2,t}$ are residuals $\hat{\epsilon}_t = y_t - \hat{\Phi} y_{t-1}$ and h_{11}, h_{12} are two bandwidths adjusted for the variation of $\hat{\epsilon}_{1,t}$ and $\hat{\epsilon}_{2,t}$, respectively. We get:

$$\begin{aligned}\hat{l}_T(y_1, y_2 | Y_T) \quad = \quad & \frac{\frac{1}{T}\frac{1}{h_2}\sum_{t=1}^{T} K\left(\frac{\hat{a}^{21}(y_1 - y_{1,t}) + \hat{a}^{22}(y_2 - y_{2,t})}{h_2}\right)}{\frac{1}{T}\frac{1}{h_2}\sum_{t=1}^{T} K\left(\frac{\hat{a}^{21}(y_{1,T} - y_{1,t}) + \hat{a}^{22}(y_{2,T} - y_{2,t})}{h_2}\right)} \\ & |\hat{\lambda}_2| \frac{1}{T}\frac{1}{h_{11}h_{12}}\sum_{t=1}^{T} K\left(\frac{y_1 - \hat{\phi}_{1,1}y_{1,T} - \hat{\phi}_{1,2}y_{2,T} - \hat{\epsilon}_{1,t})}{h_{11}}\right) \\ & K\left(\frac{y_2 - \hat{\phi}_{2,1}y_{1,T} - \hat{\phi}_{2,2}y_{2,T} - \hat{\epsilon}_{2,t})}{h_{12}}\right)\end{aligned}$$

On-Line Appendix D: Empirical Analysis of Oil Prices and GDP

1. Quarterly Data Estimation

1.1 Estimated errors

The estimated errors of the mixed VAR(1) model are plotted below:

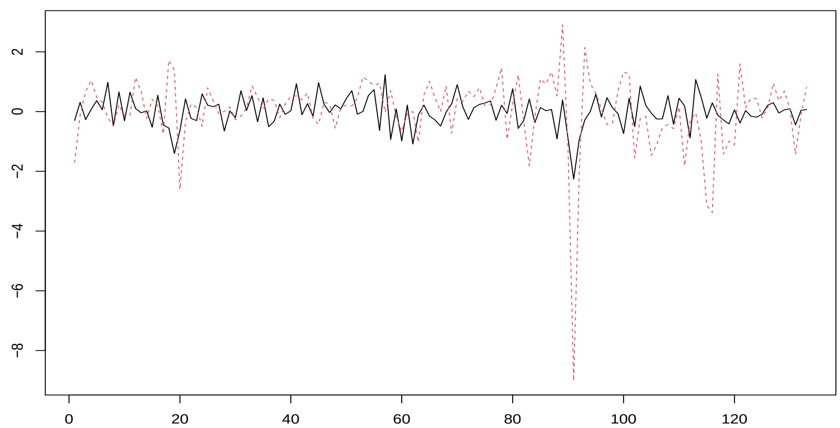
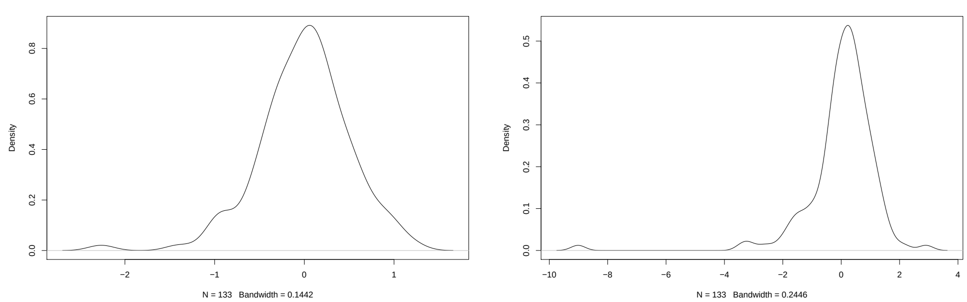


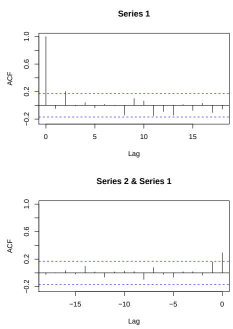
Figure 1: Errors $\hat{\epsilon}_t$, Noncausal VAR(1), Q1 1986 -Q2 2019

Their densities are non-Gaussian, as shown in Figure 2:

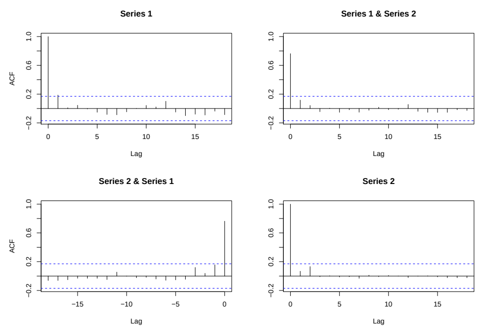


(a) Density of $\hat{\epsilon}_{1,t}$ (b) Density of $\hat{\epsilon}_{2,t}$

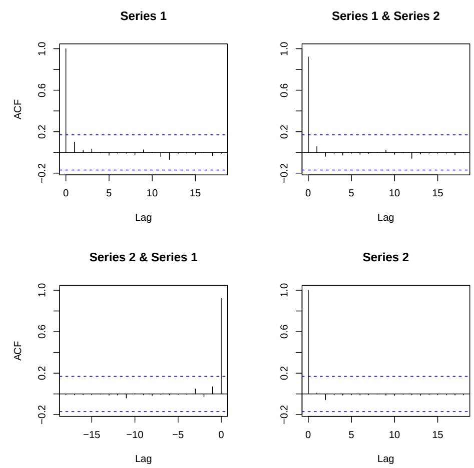
Figure 2: Density of $\hat{\epsilon}_t$



(a) ACF of $\hat{\varepsilon}$



(b) ACF of $\hat{\varepsilon}^2$



(c) ACF of $\hat{\varepsilon}^3$

Figure 3: Residual ACF

We observe that the estimated residuals, their squares and third powers are serially uncorrelated. The variance-covariance matrix is $\hat{\Sigma} = \begin{bmatrix} 0.262 & 0.184 \\ 0.184 & 1.506 \end{bmatrix}$ and contemporaneous correlation is 0.28 (statistically significant at 0.05).

1.1 GCov estimation of the mixed model for log-prices

The estimation of the VAR(1) model applied to the logarithms of oil prices produces an autoregressive matrix with two imaginary eigenvalues of modulus greater than 1. The estimation applied to the differences of log-prices produces an autoregressive matrix with two real eigenvalues inside the unit circle.

The figure below shows a path of a process with the same matrix Φ and t(6) distributed errors with the variance-covariance matrix equal to $\hat{\Sigma}$ of the mixed VAR(1) given above.

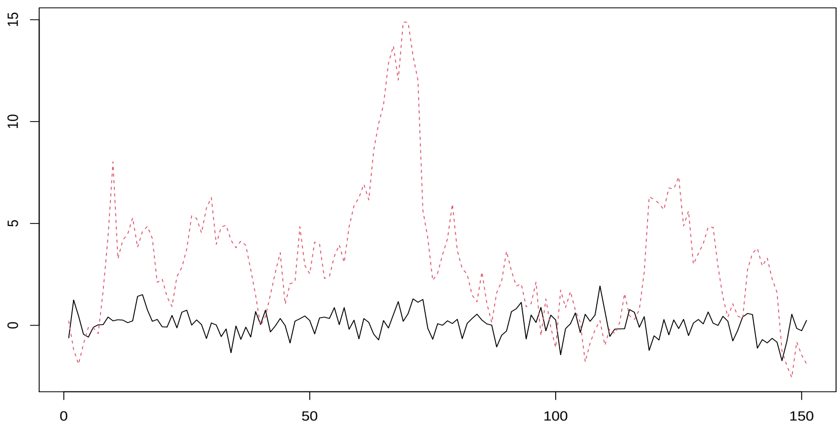


Figure 4: Simulated path

We observe that the simulated model can imitate the dynamics of a bivariate series of growth and oil.

1.3 Analysis of the SVAR model by OLS

The Augmented Dickey-Fuller test and Phillips-Perron test performed without intercept, with intercept and with both intercept and trend do not reject the unit root hypothesis in the oil prices. All these tests reject the unit root in the GDP rate (results available on request). These standard testing methods are based on the assumption of linear causal dynamics and can be misleading. In our framework of a mixed causal-noncausal model, the causal dynamic is stationary, nonlinear with local trends. The Dickey-Fuller and Phillips-Perron tests do not distinguish between the local and global trends [Gourieroux, Jasiak (2019)]. Hence their outcomes can be confusing.

The VAR(1) estimated by the OLS and Normality-based Maximum Likelihood (ML) produce similar results. The OLS output is given below:

The first dependent variable (top panel) is the oil price and the second dependent variable (bottom panel) is the GDP. The regressors $x1$ and $x2$ are the lagged oil and lagged GDP rate. The coefficient on the lagged oil is close to 1 and its ML estimator is 0.95 with standard error 0.029.

Next, we adopt the approach of Kilian, Vigfusson (2017) and model the difference of logarithms of oil prices and GDP rate as a VAR(1) model.

```
 cases:                  133      Dependent variable:                Y
Missing cases:             0          Deletion method:             None
Total SS:         161256.914          Degrees of freedom:           130
R-squared:             0.895          Rbar-squared:               0.894
Residual SS:       16905.724          Std error of est:          11.404
F(2,130):            555.009          Probability of F:           0.000
Durbin-Watson:         1.570

                        Standard                 Prob   Standardized  Cor with
Variable     Estimate    Error       t-value     >|t|     Estimate    Dep Var
-------------------------------------------------------------------------------
CONSTANT     2.535290   2.622297    0.966820     0.335      ---         ---
X1           0.950855   0.029383   32.360995     0.000    0.950777    0.945956
X2           1.151480   1.799521    0.639881     0.523    0.018800   -0.225014

Valid cases:                 133      Dependent variable:                Y
Missing cases:                 0      Deletion method:                None
Total SS:                 42.900      Degrees of freedom:              130
R-squared:                 0.211      Rbar-squared:                  0.199
Residual SS:              33.842      Std error of est:              0.510
F(2,130):                 17.396      Probability of F:              0.000
Durbin-Watson:             2.126

                        Standard                 Prob   Standardized  Cor with
Variable     Estimate    Error       t-value     >|t|     Estimate    Dep Var
-------------------------------------------------------------------------------
CONSTANT     0.731832   0.117326    6.237568     0.000      ---         ---
X1          -0.004524   0.001315   -3.440987     0.001   -0.277323   -0.354787
X2           0.301778   0.080514    3.748142     0.000    0.302078    0.373194
```

Figure 5: Results of OLS estimation of VAR(1) with price levels

Valid cases:	133	Dependent variable:	Y
Missing cases:	0	Deletion method:	None
Total SS:	0.625	Degrees of freedom:	130
R-squared:	0.046	Rbar-squared:	0.031
Residual SS:	0.596	Std error of est:	0.068
F(2,130):	3.099	Probability of F:	0.048
Durbin-Watson:	1.930		

Variable	Estimate	Standard Error	t-value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.007812	0.008969	-0.871063	0.385	---	---
X1	0.153936	0.086580	1.777968	0.078	0.155994	0.180885
X2	0.013962	0.010578	1.319963	0.189	0.115810	0.149338

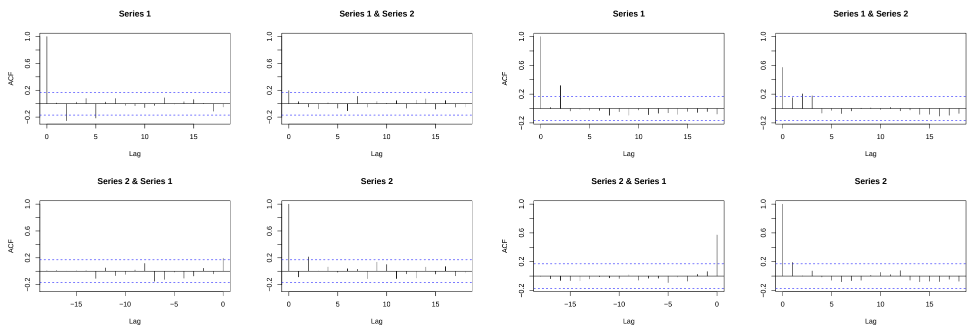
Valid cases:	133	Dependent variable:	Y
Missing cases:	0	Deletion method:	None
Total SS:	42.900	Degrees of freedom:	130
R-squared:	0.141	Rbar-squared:	0.128
Residual SS:	36.859	Std error of est:	0.532
F(2,130):	10.653	Probability of F:	0.000
Durbin-Watson:	2.166		

Variable	Estimate	Standard Error	t-value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	0.393880	0.070509	5.586242	0.000	---	---
X1	-0.328474	0.680670	-0.482575	0.630	-0.040171	0.041897
X2	0.381448	0.083159	4.586962	0.000	0.381828	0.373194

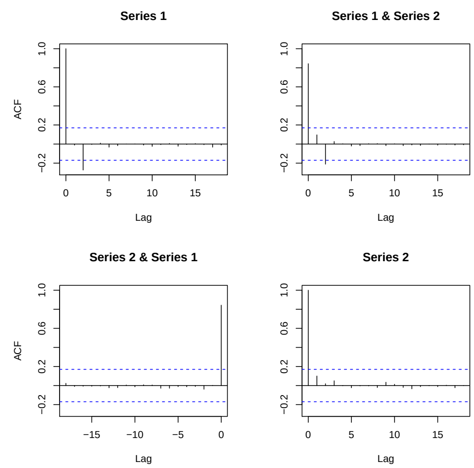
Figure 6: Results of OLS estimation of VAR(1) with diff log prices

The first regressor $x1$ is the lag 1 of differenced logarithm of oil price. The second regressor $x2$ is the lagged GDP rate.

The ACF of the $\hat{\epsilon}_t$ of this model is given below.



(a) ACF of $\hat{\epsilon}$ (b) ACF of $\hat{\epsilon}^2$



(c) ACF of $\hat{\epsilon}^3$

Figure 7: Power ACF of $\hat{\epsilon}_t$, diff-log prices, VAR(1)

We observe that the residuals of differenced logarithms of oil prices show significant autocorrelation at high lags and an ARCH effect at lag 2. There also remains significant correlation at lag 2 in powers 3 of the residuals. This motivates extending the lag up to lag 4, as in Kilian, Vigfusson (2017) and adding a comtemporaneous difference log of oil prices to the GDP rate equation as regressor $x9$, which gives this model a structural interpretation.

File Edit Format View Help

Valid cases:	130	Dependent variable:	Y
Missing cases:	0	Deletion method:	None
Total SS:	0.588	Degrees of freedom:	121
R-squared:	0.108	Rbar-squared:	0.049
Residual SS:	0.524	Std error of est:	0.066
F(8,121):	1.838	Probability of F:	0.076
Durbin-Watson:	2.012		

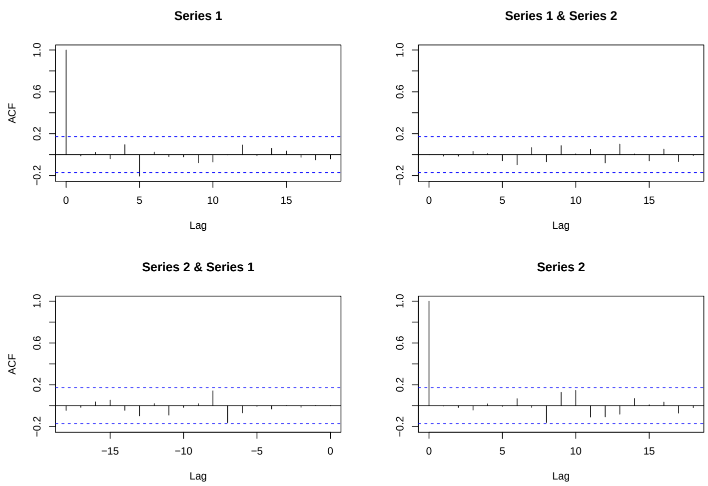
Variable	Estimate	Standard Error	t-value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.004545	0.011101	-0.409448	0.683	---	---
X1	0.166449	0.093142	1.787039	0.076	0.166437	0.143128
X2	0.017744	0.011516	1.540842	0.126	0.151291	0.160358
X3	-0.264640	0.094306	-2.806189	0.006	-0.264952	-0.221230
X4	-0.001098	0.012009	-0.091452	0.927	-0.009374	-0.002212
X5	0.086105	0.092240	0.933485	0.352	0.088086	-0.004605
X6	-0.007252	0.011947	-0.606970	0.545	-0.061769	-0.036153
X7	-0.055362	0.089032	-0.621817	0.535	-0.057529	0.005061
X8	0.001570	0.011489	0.136642	0.892	0.013385	0.007188

Valid cases:	130	Dependent variable:	Y
Missing cases:	0	Deletion method:	None
Total SS:	42.752	Degrees of freedom:	120
R-squared:	0.247	Rbar-squared:	0.190
Residual SS:	32.201	Std error of est:	0.518
F(9,120):	4.369	Probability of F:	0.000
Durbin-Watson:	2.001		

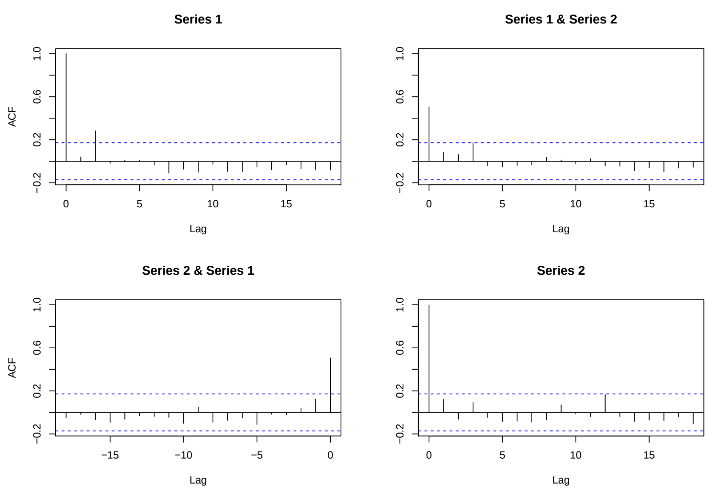
Variable	Estimate	Standard Error	t-value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	0.293411	0.087406	3.356889	0.001	---	---
X1	-0.887984	0.742475	-1.195979	0.234	-0.104150	0.049279
X2	0.272869	0.091494	2.982377	0.003	0.272897	0.378264
X3	0.183911	0.765787	0.240160	0.811	0.021598	0.039145
X4	0.243701	0.094494	2.579001	0.011	0.243989	0.334491
X5	-0.654226	0.728378	-0.898196	0.371	-0.078504	-0.049393
X6	-0.009933	0.094145	-0.105509	0.916	-0.009925	0.159095
X7	-0.864326	0.701648	-1.231852	0.220	-0.105350	-0.120046
X8	0.030606	0.090408	0.338536	0.736	0.030608	0.123363
X9	1.759205	0.715295	2.459410	0.015	0.206348	0.230292

Figure 8: Results of OLS estimation of SVAR(4) with diff-log prices

The ACF of the residuals of this SVAR model are shown below.



(a) ACF of $\hat{\epsilon}$



(b) ACF of $\hat{\epsilon}^2$

Figure 9: Power ACF of $\hat{\epsilon}_t$, diff-log prices, SVAR(4)

The residuals are contemporaneously uncorrelated. There is a significant autocorrelation in the squared oil residuals at lag 2 that was also observed in the ACF of the previous VAR(1) model. This nonlinear effect has not been removed by extending the lags in the model.

1.4 Analysis of the threshold VAR model, OLS

Next, we estimate model (1) on page 1753 of Kilian, Vigfusson (2017) by adding the terms representing the differences between the log oil prices and its maximum over the past 3 years. These terms at lags 0 to 4, appear in the model below as regressors $x11$ to $x13$.

Valid cases:	130	Dependent variable:	Y
Missing cases:	0	Deletion method:	None
Total SS:	0.588	Degrees of freedom:	121
R-squared:	0.108	Rbar-squared:	0.049
Residual SS:	0.524	Std error of est:	0.066
F(8,121):	1.838	Probability of F:	0.076
Durbin-Watson:	2.012		

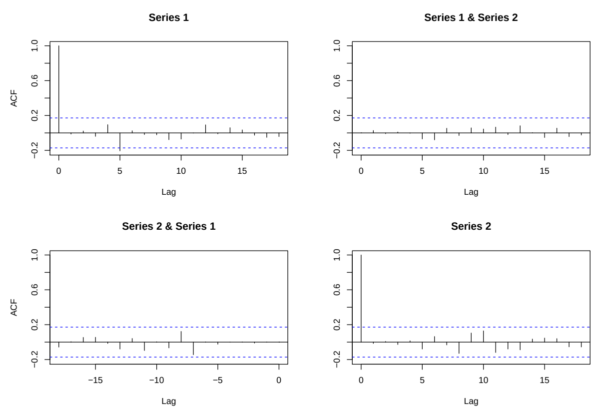
Variable	Estimate	Standard Error	t-value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	-0.004545	0.011101	-0.409448	0.683	---	---
X1	0.166449	0.093142	1.787039	0.076	0.166437	0.143128
X2	0.017744	0.011516	1.540842	0.126	0.151291	0.160358
X3	-0.264640	0.094306	-2.806189	0.006	-0.264952	-0.221230
X4	-0.001098	0.012009	-0.091452	0.927	-0.009374	-0.002212
X5	0.086105	0.092240	0.933485	0.352	0.088086	-0.004605
X6	-0.007252	0.011947	-0.606970	0.545	-0.061769	-0.036153
X7	-0.055362	0.089032	-0.621817	0.535	-0.057529	0.005061
X8	0.001570	0.011489	0.136642	0.892	0.013385	0.007188

Valid cases:	130	Dependent variable:	Y
Missing cases:	0	Deletion method:	None
Total SS:	42.752	Degrees of freedom:	115
R-squared:	0.337	Rbar-squared:	0.256
Residual SS:	28.366	Std error of est:	0.497
F(14,115):	4.166	Probability of F:	0.000
Durbin-Watson:	2.030		

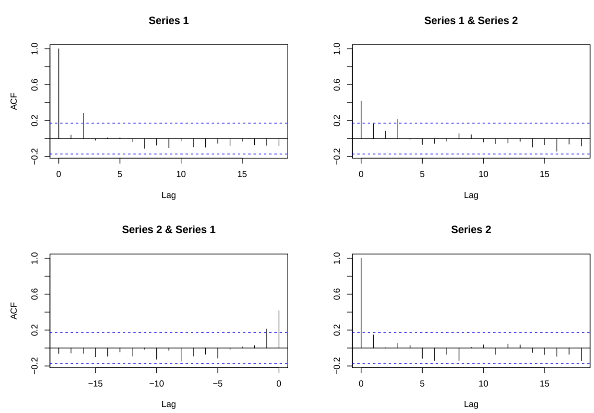
Variable	Estimate	Standard Error	t-value	Prob >|t|	Standardized Estimate	Cor with Dep Var
CONSTANT	0.339777	0.122609	2.771227	0.007	---	---
X01	0.062649	0.996067	0.062896	0.950	0.007348	0.049279
X02	0.186963	0.095677	1.954096	0.053	0.186982	0.378264
X03	0.099285	0.992394	0.100046	0.920	0.011659	0.039145
X04	0.293071	0.097102	3.018183	0.003	0.293417	0.334491
X05	-1.977873	0.904666	-2.186302	0.031	-0.237335	-0.049393
X06	0.070690	0.096619	0.731642	0.466	0.070630	0.159095
X07	-1.165629	0.884270	-1.318181	0.190	-0.142075	-0.120046
X08	0.023176	0.091962	0.252021	0.801	0.023178	0.123363
X09	3.634342	0.901947	4.029440	0.000	0.426295	0.230292
X10	-8.043220	2.422159	-3.320681	0.001	-0.371093	-0.030453
X11	-0.110458	2.492309	-0.044319	0.965	-0.005096	-0.043937
X12	1.762945	2.461134	0.716314	0.475	0.081338	-0.062059
X13	4.574759	2.286909	2.000411	0.048	0.211067	-0.025177
X14	-3.042409	2.075715	-1.465716	0.145	-0.140521	-0.175162

Figure 10: Results of OLS estimation of Kilian-Vigfusson model

The model contains 22 regressors and accounts for the simultaneity and nonlinear effects. Moreover, the model is structural and the current value of oil appears in the growth rate equation 2. The oil is modelled as the first differences in log prices. The ARCH effect in the residuals at lag 2 is not removed when the additional regressors are added to the model. This nonlinear effect remains, as shown in the residual ACF below.



(a) ACF of $\hat{\epsilon}$



(b) ACF of $\hat{\epsilon}^2$

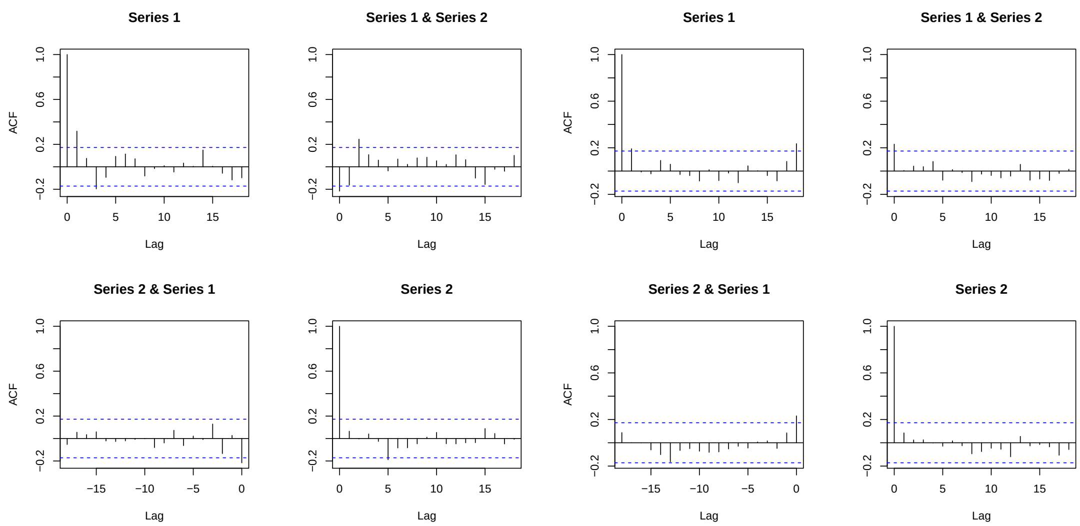
Figure 11: Power ACF of $\hat{\epsilon}_t$, OLS estimation

We conclude that the mixed causal-noncausal VAR(1) model with 4 autoregressive coefficients fits the data equally well as the above model with 24 coefficients.

It is interesting to see if the residuals of the model remain uncorrelated when the normality assumption is relaxed and the GCov estimator is used instead of the OLS estimator.

The Shapiro test of normality applied to the residuals of the oil equation does not reject the null (p-value of 0.8387). However, the normality is strongly rejected in the residuals from the second equation (p-value of 0.01711).

The autocorrelations of the residuals and their squares are plotted below:



(a) ACF of GCov residual (b) ACF of squared GCov residual

Figure 12: Power ACF of $\hat{\epsilon}_t$, GCov estimation

We observe that unlike the OLS residuals, the GCov residuals of the Kilian-Vigfusson model are contemporaneously correlated. Their squares are contemporaneously correlated too. The normality assumption and the use of OLS has an effect on the estimates.

2. Predictive Densities

We first provide the predictive densities of the estimated models of Section 6.2 for additional dates between T=112 and T=118.

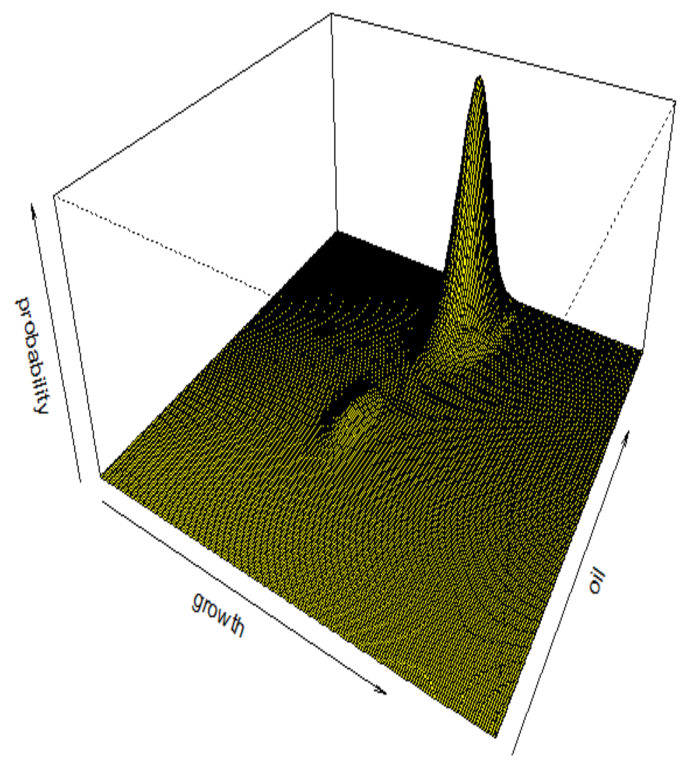


Figure 13: VAR(1) based predictive density estimation, T=112

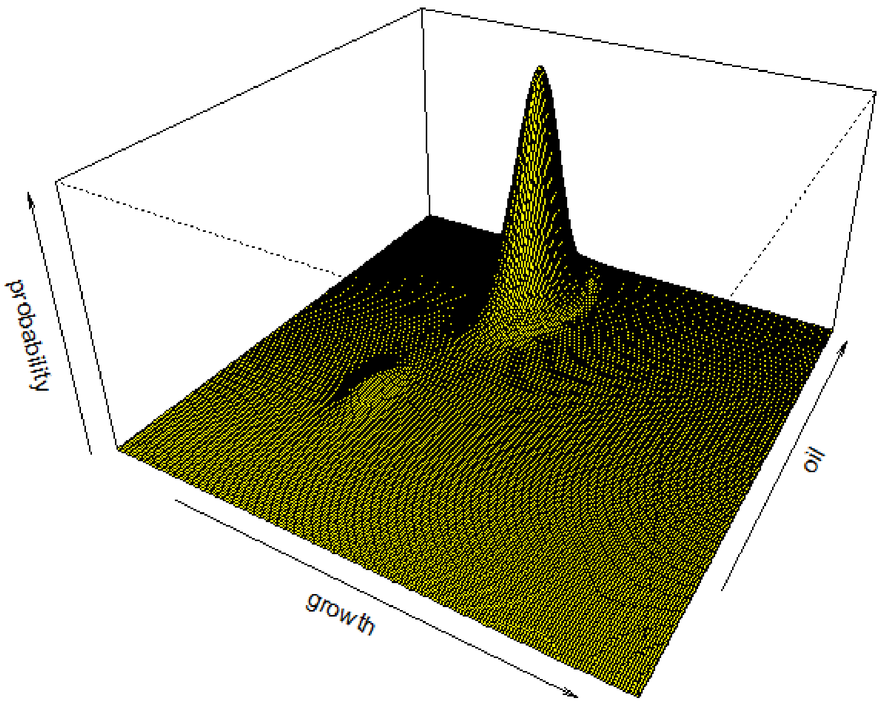


Figure 14: VAR(1) based predictive density estimation, T=114

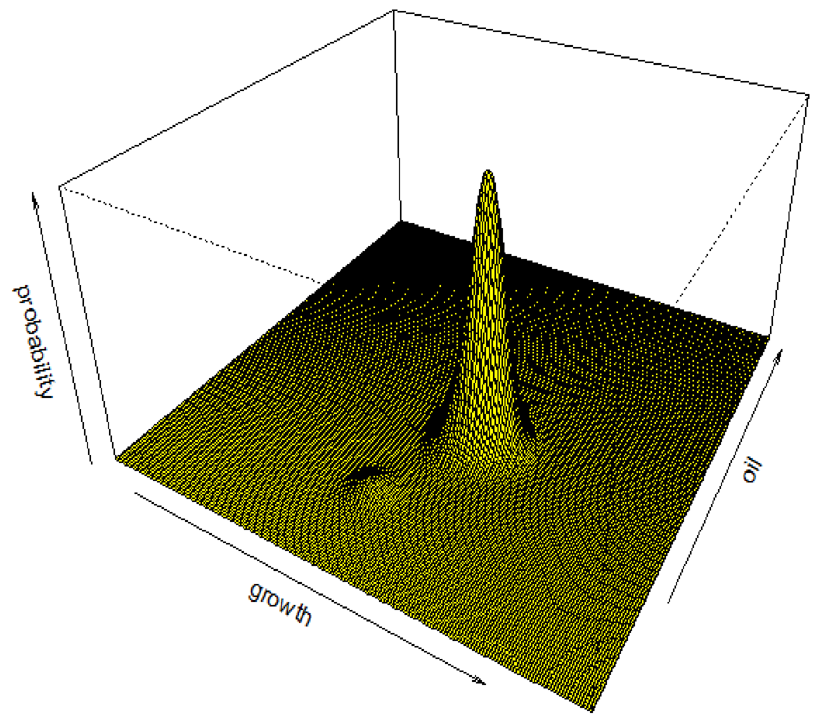


Figure 15: VAR(1) based predictive density estimation, T=115

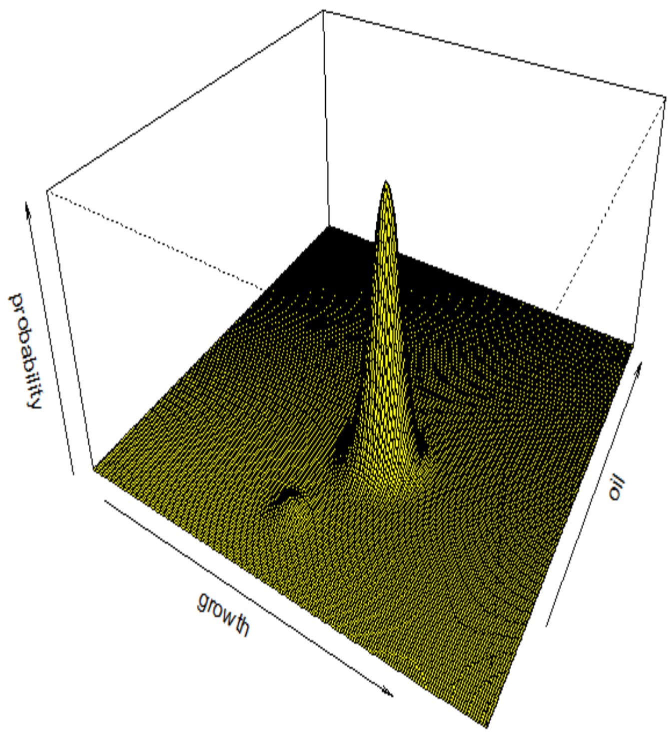


Figure 16: VAR(1) based predictive density estimation, T=116

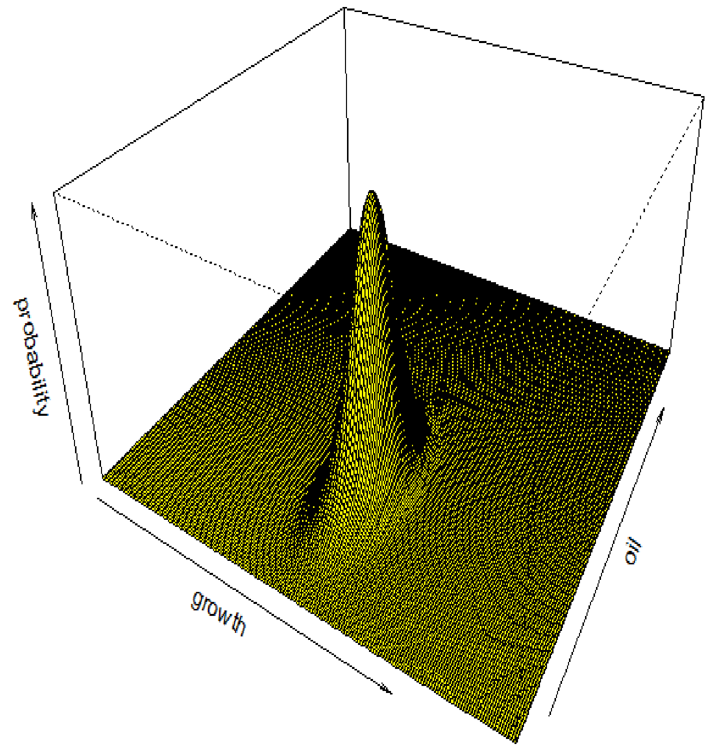


Figure 17: VAR(1) based predictive density estimation, T=118

2.2. One step-ahead path forecasting

We perform a sequence of 1 step ahead forecasts, where the series is re-estimated at each step on a subsample of observations from time 1 up to the forecast origin. The first forecast is performed at time T=100 and the values of both component series at time T+1=101 are predicted. Next, the sample is increased by 1 observation and extended up to time T=101. Then the value of the series at time T+1=102 is forecast, and so on, until the end of the path.

At each step, the estimated parameters are the sample means and the parameters of matrix Φ, which are used in the forecast computation. The forecast is evaluated each time over a grid of 200 values equally spaced by 0.1, below and above the last values of latent components of the series.

The first sequence of forecasts is based on the estimates from 4 nonlinear transforms in the objective function of the GCov estimator. Only the powers 1 and 2 of the errors are considered. The forecasts (black lines) and the true values of the processes (red lines) are given below.

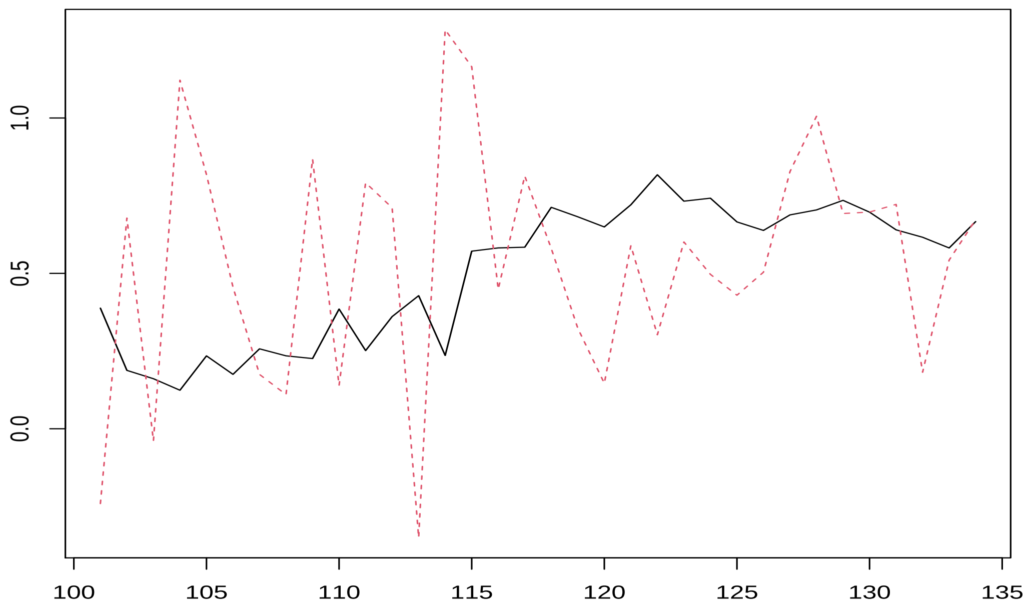


Figure 18: VAR(1) based path of growth rate (4 transforms), T=101-135

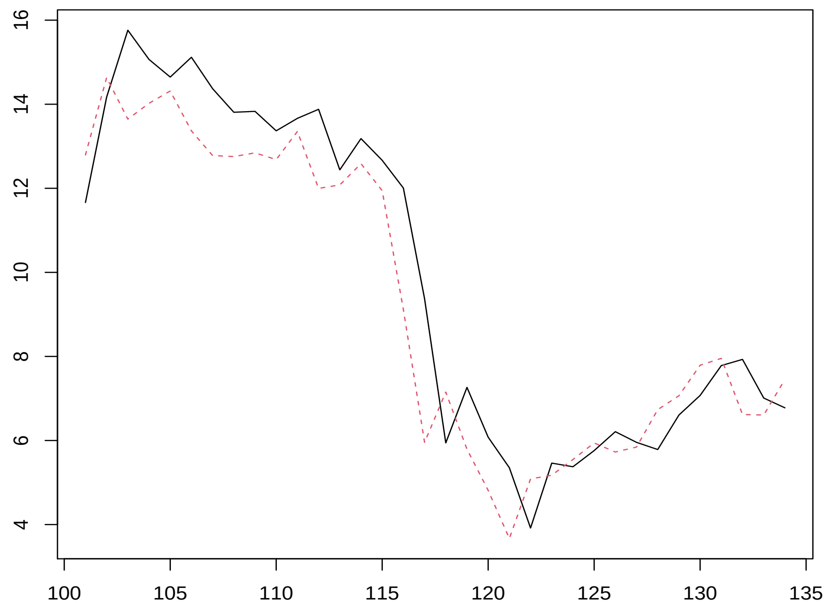


Figure 19: VAR(1) based path of oil/10 (4 transforms), T=101-135

We observe that the path of oil prices is well fitted. However, the VAR(1) has a difficulty fitting small variations in the growth rate. The MSFE (Mean squared Forecast Error) is 0.184 for the rate and 1.582 for the oil.

For this reason, we increase the number of power transforms in the objective function, by considering powers 1, 2, 3 and 4 of the errors. The Figures below show that this refined estimation approach helps to forecast the small changes in the growth rate.

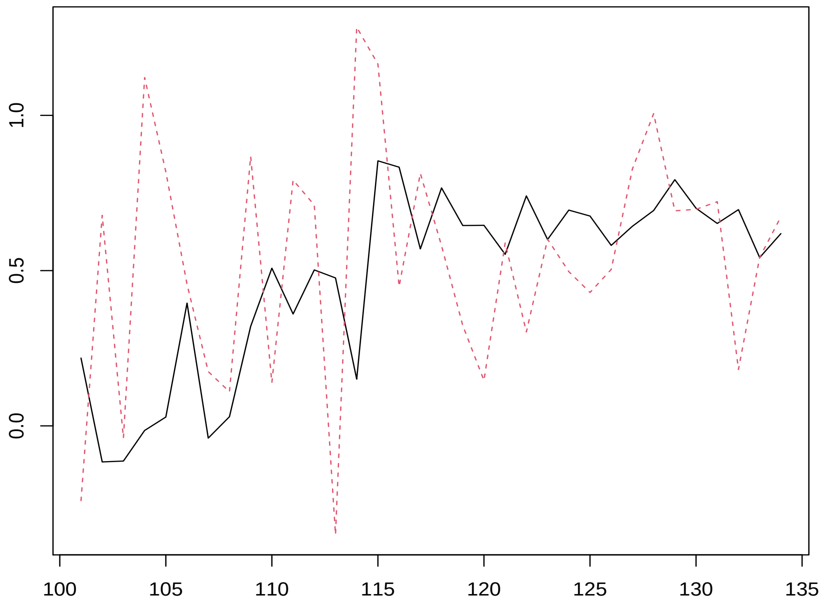


Figure 20: VAR(1) based path of growth rate (8 transforms), T=101-135

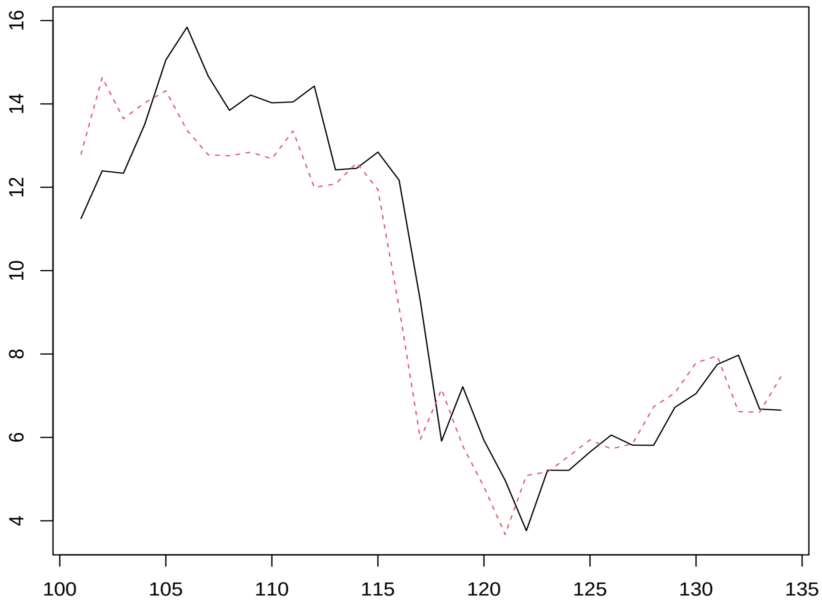


Figure 21: VAR(1) based path of oil/10 (8 transforms), T=101-135

We observe that it takes a few steps for this method to start improving the forecast of the series. Over the entire subset of 34 forecasts, the MSFE is higher under this approach, as compared to the previous one. More specifically, the MSFE of growth rate is 0.201. The MSFE of oil is 1.904. An improvement appears over the set of last 14 observations.

Over the last 14 observations, when 4 additional power functions are added to the GCov estimation, the MSFE of growth rate decreases from 0.071 to 0.068 and the MSFE of oil decreases from 0.727 to 0.645

For comparison, we compute the IN-SAMPLE fitted values of the structural model of Kilian, Vigfusson model with 22 parameters estimated by the OLS. It is specified for the differences of logarithmic values of oil and accounts for the nonlinearity in oil series and simultaneity in the growth rate. As mentioned earlier,the current value of the oil variable determines the growth rate simultaneously et each point in time (regressor 9). In our computation, the true values of oil, rather than the predicted ones are used to compute the fitted values of growth. Moreover, the levels of oil prices are computed by adding the fitted difference to the true past value of the oil price levels rather than from the random walk representation involving the initial value.

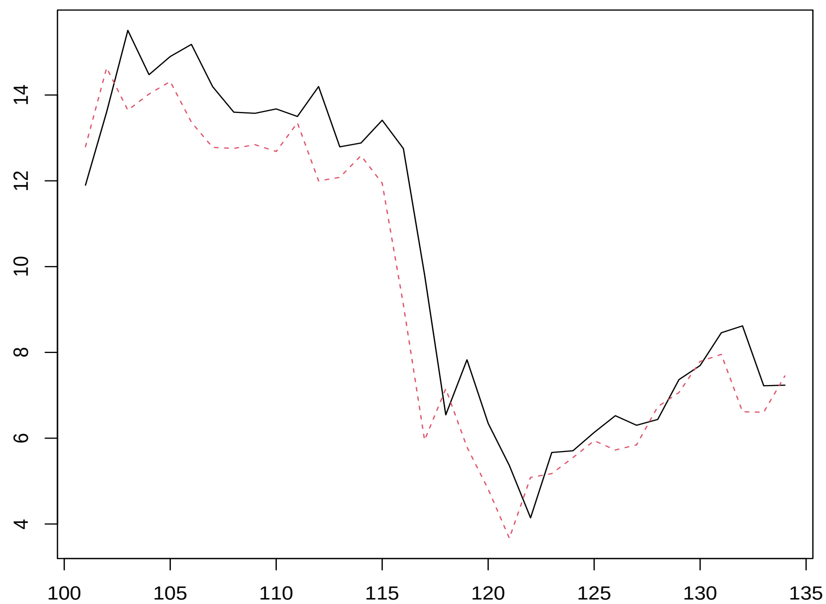


Figure 22: Kilian-Vigfusson: fitted values of oil/10), T=101-135

The MSE of fitted values of oil over the same period of 34 forecasts is 1.934 and is higher than the oos MSFE of noncausal VAR(1). Over the subset of last 14 observations, the MSE of oil is 0.867 and is also higher than the mean forecast error of the VAR(1) model.

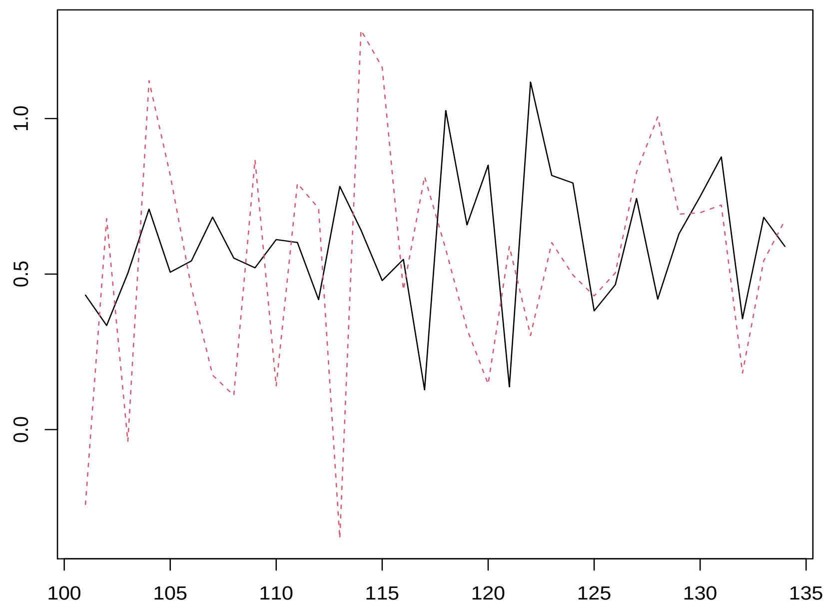


Figure 23: Kilian-Vigfusson: fitted values of growth, T=101-135

The MSE of the fitted values of growth rate is 0.204 and is marginally higher/equal, as compared to the oos MSFE of noncausal VAR(1). Over the subset of last 14 observations, the MSE of growth rate is 0.140 and is higher than the mean forecast error of the VAR(1) model. Importantly, we observe that the fitted series does not accommodate local changes in the trend of the rate.

In the table below the forecasts are computed for the levels of the variables (oil divided by 10), with the mean added to the forecast. The previous tables were forecasting the deviations from the mean.

Forecast with 8 power transforms

T	forecast		interval growth		interval oil		true values	
	growth	oil	q11	q12	q21	q22	growth	oil
110	0.507	14.025	-11.192	1.407	-2.174	15.825840	0.139	12.684
112	0.502	14.427	-10.997	1.402	-2.872	16.027	0.711	11.998
114	0.150	12.455	-7.849	1.050	-3.144	14.055	1.284	12.586
115	0.853	12.844	-11.346	2.053	-3.155	14.444	1.164	11.944
116	0.833	12.169	-11.066	1.933	-2.630	13.869	0.448	9.114
118	0.766	5.911	-9.633	1.666	-0.188	7.711	0.580	7.152
133	0.542	6.679	-8.457	1.442	-2.320	8.479	0.542	6.606
134	0.619	6.653	-9.180	1.519	-3.546	8.453	0.673	7.452

Forecast with 4 power transforms

T	forecast		interval growth		interval oil		true values	
	growth	oil	q11	q12	q21	q22	growth	oil
110	0.385	13.367	-11.814	1.285	-2.432	15.167	0.139	12.684
112	0.360	13.879	-11.639	1.260	-3.020	15.579	0.711	11.998
114	0.235	13.182	-6.764	1.135	-3.817	14.882	1.284	12.586
115	0.571	12.665	-13.128	1.471	-2.234	14.565	1.164	11.944
116	0.581	12.005	-12.518	1.481	-2.394	13.905	0.448	9.114
118	0.712	5.942	-9.887	1.712	0.042	7.842	0.580	7.152
133	0.581	7.010	-7.618	1.481	-2.589	8.810	0.542	6.606
134	0.666	6.777	-9.033	1.566	-3.322	8.577	0.673	7.452

References to On-Line Appendix D

Gourieroux, Hencic, Jasiak (2020): "Forecast Performance and Bubble Analysis in Noncausal MAR(1,1) Processes", Journal of Forecasting, 40, 301-326.

Gourieroux, Jasiak (2016): " Filtering, Prediction and Simulation Methods for Noncausal Processes", Journal of Time Series Analysis, Vol 37, 405-430.

Hecq, A. and E. Voisin (2021): "Forecasting Bubbles with Mixed Causal-Noncausal Autoregressive Models", Econometrics and Statistics, 20, 29-45.

Kilian, L. and R. J. Vigfusson (2017): "The Role of US Oil Price Shocks in Causing US Recessions", Journal of Money, Credit and Banking, 40, 1747- 1776.