Termpaper Makroökonometrie II

Bernd Prostmaier (Matrikelnummer: 12004113)

Thursday, February 03, 2022

Abstract

This termpaper tries some applications of state space modelling to macroeconomic questions. To be more concise, I tried to estimate crude oil price and log returns using both the R code from lecture macroeconometrics II and the stochvol package. Overall, it can be said that due to the high flexibility of the state-space approach, it is both a useful tool for research purposes and highly useful in addressing finance issues. It allows formulating rather complicated problems in a simple manner.

Contents

1 Introduction

1.1 FRED Data

The dataset used is provided by the Federal Reserve Bank of St. Louis and contains time series data retroactive to January 01, 1959. For the purposes of this paper, the most recent dataset, which includes observations through December 31, 2021, was used. Updated forms of the dataset can be accessed continuously **here**.

For a detailed documentation of the dataset, see **FRED-MD: A Monthly Database for Macroeconomic Research** in Journal of Business & Economic Statistics (2015).

1.2 Crude oil time series

As in my previous termpaper of macroeconometrics I, I will focus on time series data of crude oil prices. In chapter 3 of this termpaper, I will estimate both the crude oil price and the demeaned crude oil log returns.

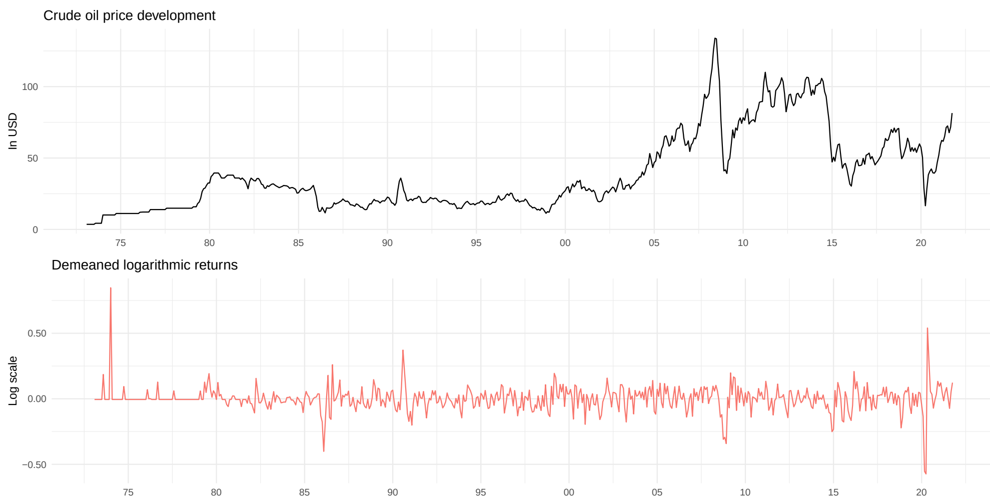

Figure 1: Vizualization of crude oil price development and demeaned log returns

2 State space modelling

Before presenting the implementation and the results of my termpaper, I will briefly present the general structure of state-space models and the Kalman filter. This section introduces general notations as well as general assumptions and their implications, which are necessary for further understanding. Overall, state space models tackle mainly two problems in macroeconomics. At first, state-space models address problems with variables that are unobservable and secondly, it addresses variables containing coefficients that are inherently time-varying, making economic relationships potentially unstable (Harvey, 1987). As we have seen in the lecture, state space models are straightforward to implement and provide simple representations of rather complex problems.

2.1 State-space model

As stated before, I will briefly present the general characteristics of state-space models and the Kalman filter. A standard state space model formulation can be represented as follows, where

$$Y_t = Z_t A_t + \epsilon_t$$

is the measurement equation (Harvey, 1987). Y_t is a vector of measured variables of dimension $n \times t$, A_t is the state vector of unobserved variables of dimension $p \times 1$. Z_t is a matrix of parameters of dimension $n \times p$ and $\epsilon_t \sim \mathcal{N}(0, H_t)$. The state equation, which describes assumptions made for latent variable A_t, is given as:

$$A_t = T_t A_{t-1} + \eta_t$$

where T_t is a matrix of parameters and $\eta_t \sim \mathcal{N}(0, Q_t)$. Q_t and H_t are referred to as hyper parameters to distinguish from other parameters (Harvey, 1987). The specification of the state space system is completed by two further assumptions: (i) the inital vector A_0 has a mean a_0 and covariance matrix P_0 and (ii) the disturbances ϵ_t and η_t are uncorrelated with each other in all time periods and with the initial state (Harvey, 1987).

Overall, the objective of state space modeling is to compute the optimal estimate of the hidden state given the observed data, which can be derived as a recursive form of Bayes's rule (Brown et al., 1998; Chen, Barbieri and Brown, 2010).

2.2 Kalman filter

To compute the likelihood of the measurement equation, one has to integrate A_t out of the complete data likelihood. To solve the integral we use the so called Kalman filter. Based on the observations up to Y_t, let a_t be the optimal estimator of A_t, P_t its covariance matrix, $a_{t|t-1}$ the estimator based on the information available in $t-1$ and $P_{t|t-1}$ its covariance matrix. The Kalman filter mainly consist of two steps. At first we compute predictions. The predicted estimate of A_t is defined by:

$$a_{t|t-1} = T_t a_{t-1}$$

with a covariance matrix $P_{t|t-1} = T_t P_{t-1} T_t' + Q_t$ (Harvey, 1989). The filtered estimate of A_t is a_t and is updated from $a_{t|t-1}$ when Y_t is known. The filtered estimate is defined as:

$$a_t = a_{t|t-1} + P_{t|t-1} Z' t F_t^{-1} (Y_t - Z_t a_{t|t-1})$$

where $F_t = Z_t P_{t|t-1} Z_t' + H_t$ and $P_t = P_{t|t-1} - P_{t|t-1} Z_t' F_t^{-1} Z_t P_{t|t-1}$. Those three relations are known as the Kalman filter equations (Anderson and Moore, 1979; Harvey, 1981). Finally the smoothed estimate of A_t is $a_{t|T}$ and is updated from a_t using the whole set of information available. They are calculated working backwards from the last value of the filtered estimate (Ducan and Horn 1972; Harvey, 1981). The Kalman filter can be viewed as a Bayesian approach to economic modelling - it consists in choosing a_t so that it maximises the probability that the observed Y_t would take place (Harvey 1989), i.e. a_t is a solution of $Max\, p(A_t|Y_t) = p(A_t, Y_t)/p(Y_t)$.

As stated by Crowder 1976 and Schweppe 1965, the likelihood function for a gaussian model can be written as:

$$Ln\ell = -\frac{N(t-k)}{2} Ln(2\pi) - \frac{1}{2} \sum_{t=k}^{T} Ln|Z_t P_{t|t-1} Z_t' + H_t| - \frac{1}{2} \sum_{t=k}^{T} v_t' (Z_t P_{t|t-1} - Z_t' + H_t)^{-1} v_t$$

where N is the number of elements contained in the vector Y_t, k the number of periods needed to derive estimates of the state vector and v_t as the vector of prediction errors $v_t = Y_t - Z_t a_{t|t-1}$. Performing all the steps above, one can see that the likelihood function can now be expressed as a function of the one-step-ahead prediction errors. The Kalman filter will then allow us to estimate the unobserved variable A_t.

2.3 Assumptions

The preceding procedure presupposes assumptions relating (i) the measurement and transition equation distrubances, (ii) initial conditions and (iii) linearity. This assumptions will now be briefly discussed.

If the measurement and transition equation disturbances are correlated, the Kalman filter needs to be modified. Jazwinksi (1970) provides answers to dealing with correlation between the two disturbances. Another potential problem is resulting by the dependency of the posterior estimates on prior information. The major constraint of the Kalman filter is that it imposes a linear structure. However, as we have seen during the lecture, if the relations of interest are believed to be non-linear, there are some potential solutions to overcome this problem, e.g. approximation/linearization, using particle filter and stochastic volatility models.

3 Estimating crude oil price and demeaned log returns

3.1 Lecture code

For estimating our state space model, I will first use the code we have seen in the lecture, using the sampler from *stochvol* package built in our custom MCMC algorithm. Here we access the package to draw from the posterior distribution of the coefficients in the state equation of volatility and to draw from the latent volatilities. At this point, I will briefly discuss the most relevant preliminaries used for estimation, whereas the whole code can be seen in the .rmd file. Firstly, we have to set the prior of initial state for initializing the Kalman filter. The first value comes from a normal distribution centered on 0 with variance of 10000, which is quite uninformative. The number of draws discarded as burn-in equals 200 (nsave = 200). After preparing the data vector y (crude oil prices), I initialized the stochastic volatility quantities: (i) mean value of autoregressive process equals 0, (ii) AR1 parameter equals 0.9 and (iii) the error variance equals 0.1. For our first estimation we set the prior for the measurement error variance, which indicates how large variances in volatility equations can be, to 1. If this prior were 0, we would pull a lot of mass toward 0 and error variances therefore would be very close to zero. After setting priors for measurement error variance, for state equation and measurement error variance, the model starts the MCMC loop.

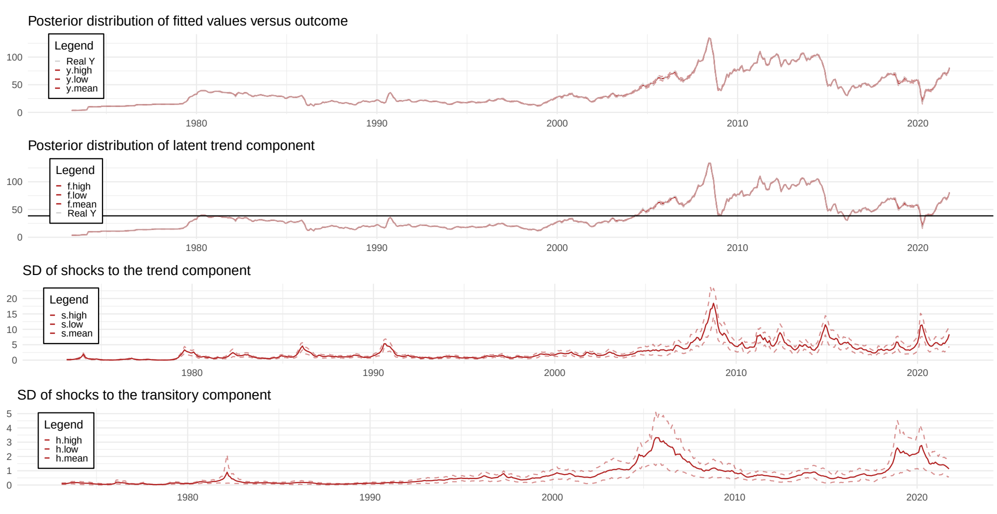

Figure 2: Estimation 1: Results estimating crude oil price

We can see in the first plot of figure 2, that the posterior intervals reproduce the real data very well (see posterior distribution of fitted values versus outcome). Furthermore, we see that the posterior distribution of latent trend component looks very similar to the posterior distribution of fitted values, indicating potentially overfitting problems, resulting from too large variances (see plot three and four in figure 2). Plot three in figure 2 shows increasing variance to the trend component in times of financial crisis and during COVID-19 pandemic. Measurement errors also significantly increases during the same periods.

In order to overcome potentially overfitting problems, I tried to set the prior mean for measurement errors very low (mu = -10, sd = 0.001). I also decreased the prior for measurement error variance (= 1e-10), which, as I already said, controls the variance within volatility equations. Figure 3 impressively shows the effect of changing some prior information. Our estimate now outsources a lot of weight.

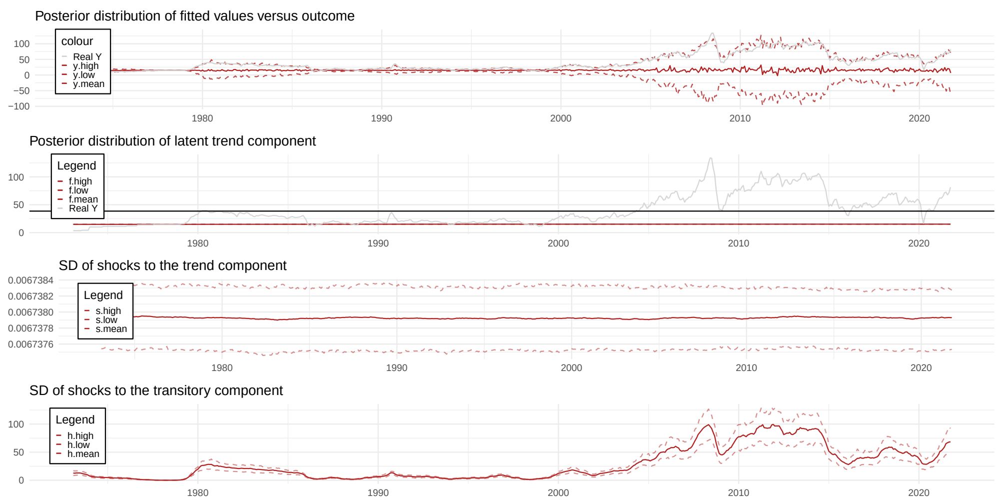

Figure 3: Estimation 2: Changing prior information

In this scenario, the latent trend is no longer a real trend, but rather resembles a constant. Standard deviation of shocks to the trend component now decreased significantly! However, we see that in times of increasing variance, the confidence intervals diverge significantly. The true value of y is no longer matched exactly, which does not mean that the model does not work, but rather is evidence that the change in prior information corresponds to the expected effects.

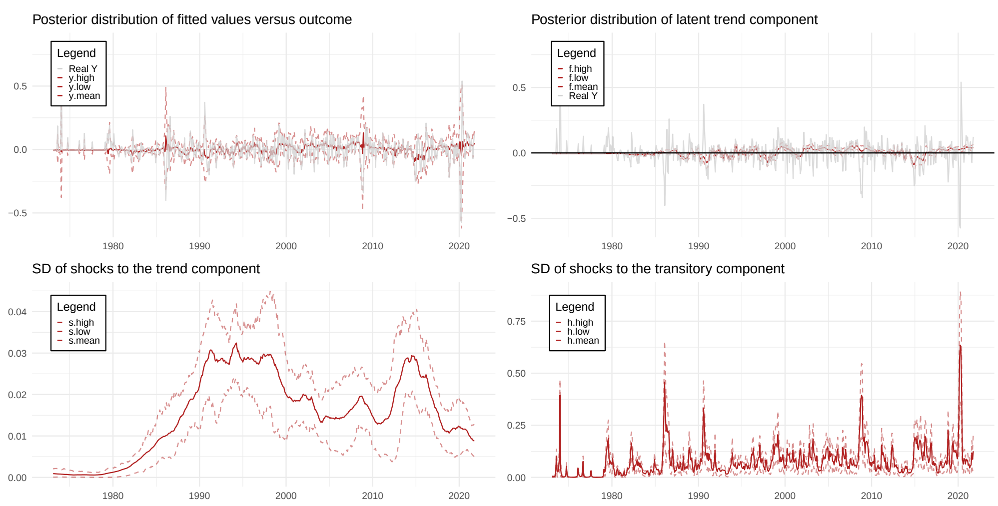

Figure 4: Estimation 3: Results estimating demeaned crude oil log returns

Beside estimating the crude oil price itself, we can also try to estimate the crude oils demeaned log returns, which have a much higher volatility. Using the set of parameter from our first estimation above, one receives quite an interesting model fit which can be seen in picture 4. Although the posterior distribution of fitted values versus outcome is quite good in my eyes, the error variance of the latent trend is probably too high.

This may be due to the fact that too few draws were made. In the following chapter I will try to estimate the log returns again by using the standalone version of the stochvol package and calculating over 10,000 iterations.

3.2 Stochvol package

Beside using *stochvol* in combination with our own MCMC algorithm, *stochvol* can be used on a stand alone basis. It provides a fully Bayesian implemention of heteoskedasticity modeling within the framework of stochastic volatility. To conduct inference, it utilizes Markov chain Monte Carlo (MCMC) samplers by obtaining draws from the posterior distribution of parameters and latent variables. For details concerning the algorithm please see the paper by Kastner and Frühwirth-Schnatter (2014) and Hosszejni and Kastner (2019).

svsample simulates from the joint posterior distribution of the SV parameters, along with the latent log variables and returns the MCMC draws. The function allows the following main arguments: (i) *priormu* indicating mean and standard deviation for the Gaussian prior distribution, (ii) *priorphi* indicating the shape parameters for the Beta prior distribution of the transformed parameter $(\phi + 1)/2$, where phi denotes the persistence of the log-volatility and (iii) *priorsigma* indicating the scaling of the transformed parameter σ^2, where sigma denotes the volatility of log-volatility.

In the following example we will estimate demeaned crude oil log returns again and set: (i) a practically uninformative *priormu* of (0, 100), which is according to Kastner (2019) a common strategy because the likelihood usually carries enough information about this parameter, (ii) a *priorphi* of (5, 1.5), which constitutes a prior that puts some belief in a persistent log-volatility but also encompasses the region where phi is around 0 and (iii) a *priorsigma* of 1 which constitutes a reasonably vague prior (according to Kastner, 2019, it should not be set too small unless there is a very good reason, e.g. explicit prior knowledge). Other method arguments can be seen at the package documentation.

```
res = svsample(FRED2$Ret, priormu = c(0,1000), priorphi = c(5, 1.5), priorsigma = 1,
               burnin = 1000, draws = 10000)
summary(res)
```

```
##
## Summary of 'svdraws' object
##
## Prior distributions:
## mu        ~ Normal(mean = 0, sd = 1000)
## (phi+1)/2 ~ Beta(a = 5, b = 1.5)
## sigma^2   ~ Gamma(shape = 0.5, rate = 0.5)
## nu        ~ Infinity
## rho       ~ Constant(value = 0)
##
## Stored 10000 MCMC draws after a burn-in of 1000.
## No thinning.
##
## Posterior draws of SV parameters (thinning = 1):
##              mean     sd     5%    50%    95%  ESS
## mu         -5.789 0.2654 -6.223 -5.787 -5.354 4186
## phi         0.858 0.0344  0.798  0.861  0.910  586
## sigma       0.826 0.0930  0.683  0.821  0.983  359
## exp(mu/2)   0.056 0.0074  0.045  0.055  0.069 4186
## sigma^2     0.691 0.1571  0.466  0.674  0.967  359
```

The return value of the *svsample*-object is an object of type *svdraws*, which is basically a list containing: (i) the paremter draws, (2) the latent log-volatilities, (3) the initial latent log-volatility draw, (4) the data provided, (5) the sampling runtime, (6) the prior hyperparameters, (7) the thinning values and (8) summary statistics, which delivers both information about the prior distributions and the posterior draws of the sv parameters (see above).

To get a better overview of the results from our model, there is a wide variety to plot the results. Figure 5 shows a visualization of estimated contemporaneous volatilities of demeaned crude oil log returns over time. This illustrates the advantages of state-space models very impressively, since, compared to a homoscedastic model, the variance varies over time. The plot shows medians and 1%/10%/90%/99% quantiles of the posterior distribution of the volatilities over time as well as predictive distributions of future volatilities.

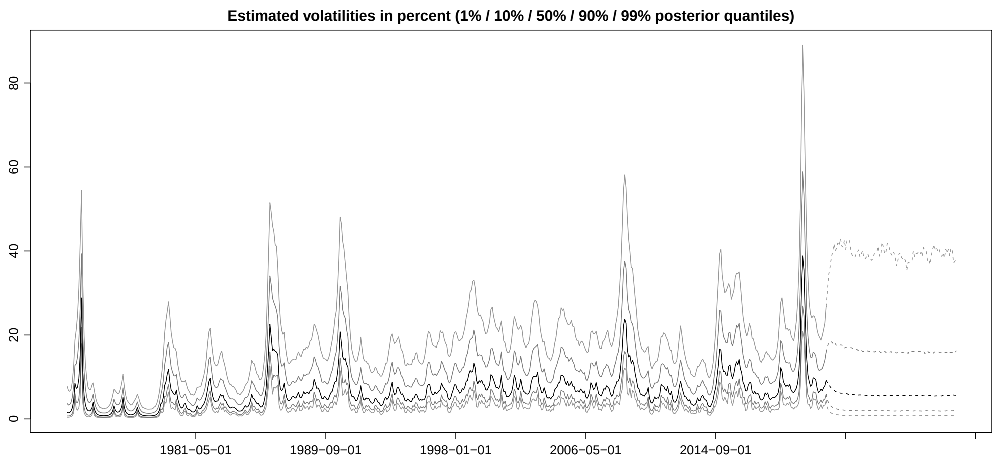

Figure 5: Estimated volatilities

Picture 6 shows trace plots for the pre specified parameters we used to initialize the model. Note that the burn-in has already been discarded.

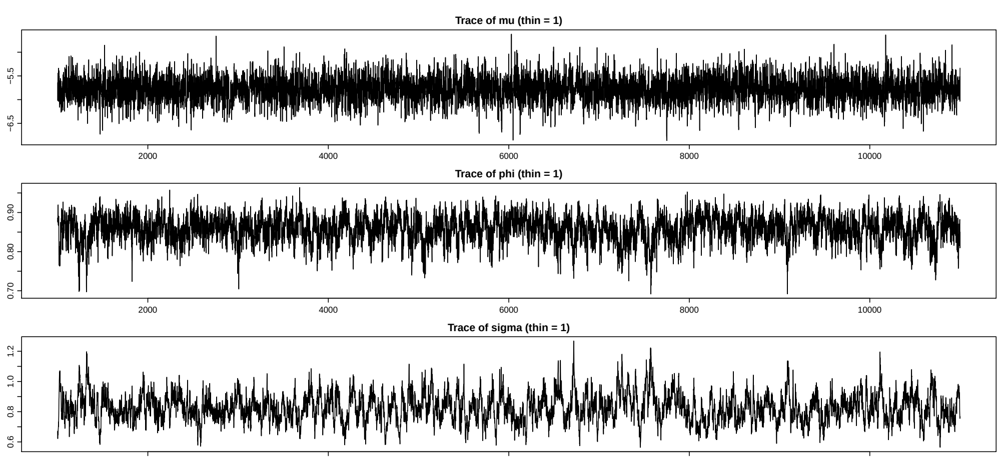

Figure 6: Traceplots of posterior draws for the model parameters

Furthermore, figure 7 shows a mean standardized residual plot for assessing the model fit. The package allows for extracting standardized residuals in a very easy way. Beside the QQ-plot one can see scatter plots with standardized residuals for each point in time. The dashed lines in the bottom left panel indicate the 2.5%/97.5% quantiles of the standard normal distribution.

One can see that using the standard parameter proposed by Kastner, 2019, and increasing the number of iterations, significantly improves the result from our state-space model. The QQ-plot shows a quite good fit,

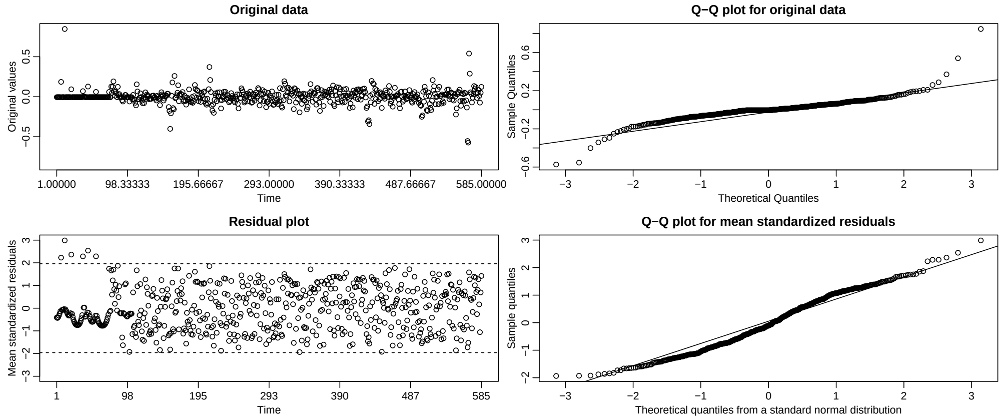

Figure 7: Estimated volatilities

which is far better than the estimate of my previous term paper using Gibbs sampling and SSVS algorithm. Figure 8 shows a nearly perfect match from our model. However, maybe the most relevant stumbling block deals with the question what our priors are as well as addressing their relevance to prevent from overfitting. The issue of relevance of priors can be addressed using various tests available (Kastner, 1989), which are beyond of this termpaper. At the end I would like to mention that even computationally intensive operations are possible in a very easy way using the stochvol package, potentially improving our estimation even further. An example of estimating parallel chaines with a pre-defined cluster object can be seen in the markdown file executing the last code chunk.

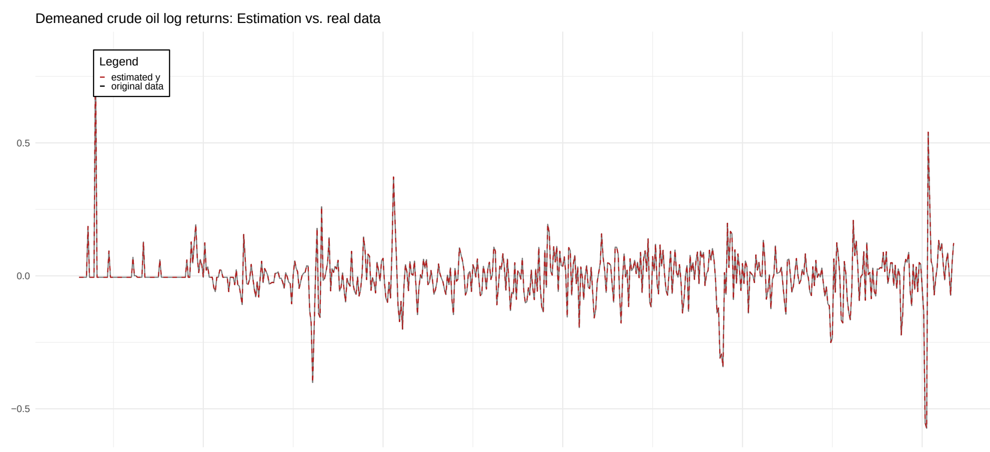

Figure 8: Model fit: Estimated data vs. real data

4 Conclusion

As we already discussed during the lecture, a key variable in the estimation of state-space models is the relative smoothness of the unobserved variable, which is governed by the relative size of the error variances in measurement and state equations. Compared to the model that I have estimated in my termpaper in macroeconometrics I, state-space models can improve the estimation of crude oil prices significantly, due to the fact that variance varies over time. In general, models that can be put into state space form allows formulating rather complicated problems in a simple manner. However, as seen in chapter 3, specifying the prior knowledge available is crucial.

5 References

Anderson, B. & Moore, J. (2005). Optimal filtering. Mineola, N.Y: Dover Publications.

Duncan, D. B., & Horn, S. D. (1972). Linear Dynamic Recursive Estimation from the Viewpoint of Regresion Analysis. Journal of the American Statistical Association, 67(340), 815–821. https://doi.org/10.2307/2284643

Harvey, A. (1981). Time series models. Cambridge, Mass: MIT Press.

Harvey, A. (1987). Applications of the Kalman filter in econometrics. In T. Bewley (Ed.), Advances in Econometrics: Fifth World Congress. Cambridge University Press.

Harvey, A. (1989). Forecasting, Structural Time Series Models and the Kalman Filter.

Kastner, G. (2016). Dealing with Stochastic Volatility in Time Series Using the R Package stochvol. Journal of Statistical Software, 69(5), 1–30. https://doi.org/10.18637/jss.v069.i05

Koop, Gary and Korobilis, Dimitris, (2009), Bayesian Multivariate Time Series Methods for Empirical Macroeconomics, Working Paper series, Rimini Centre for Economic Analysis.

Zhe Chen and Emery N. Brown (2013) State space model. Scholarpedia, 8(3):30868.