

VARsignR: Estimating VARs using sign restrictions in R

by Christian Danne

Abstract VARsignR identifies structural shocks in Vector Autoregressions (VARs) using sign restrictions. It implements Uhlig's (2005) rejection method, Uhlig's (2005) penalty function approach, the Rubio-Ramirez et al. (2010) rejection method, and Fry and Pagan's (2011) median target method. This vignette shows the usage and provides some technical information on the procedures that should help users to bridge the gap between VARsignR and the underlying technical papers.

Keywords: Sign restrictions, vector autoregression, Bayesian.

1. Introduction

Vector autoregressions (VARs) have become the cornerstone of macroeconomic policy analysis, forecasting, and testing dynamic stochastic general equilibrium (DSGE) models (Del Negro and Schorfheide, 2011). Recovering the underlying structural shocks of the model, however, has been an unresolved theoretical issue. Since VAR models are reduced form models it is impossible to structurally interpret the dynamics induced by the shocks without imposing additional restrictions (Canova, 2007). The standard approach has been to identify the structural shocks recursively using a Cholesky decomposition, or applying short-run or long-run zero restrictions (cf. Sims, 1980; Blanchard and Quah, 1989; Gali, 1992).¹ While in some cases these restrictions can be justified by economic theory through, for example, information lags or nominal rigidities, recursive structures and zero restrictions are inconsistent with most theoretical models (Canova and Pina, 2005).

Faust (1998), Canova and De Nicrolo (2002), and Uhlig (2005) show that structural inference in VAR models can be based on prior beliefs about the signs of the impacts of certain shocks derived from theoretical models. While DSGE models do not produce any zero restrictions or recursive structures, these models produce a large number of non-parametric sign restrictions that can be used for the identification process (Canova, 2007).

This paper introduces the VARsignR package containing a number of routines to recover structural shocks from VARs using sign restrictions. It currently implements the Uhlig's (2005) rejection method, Uhlig's (2005) penalty function method, Rubio-Ramirez et al. (2010, RWZ from hereafter) rejection method, and Fry and Pagan's (2011) median target (MT) method. I discuss the implementation and demonstrate the usage of the routines. Further, the paper discusses some of the shortcomings of sign restrictions, the researcher should be aware of when using VARsignR.

The remainder of the paper is as follows. Section 2 describes the general idea and implementation of sign restrictions. Section 3 discusses the use of the three core procedures of the VARsignR package. Section 4 discusses some of the problems associated with sign restrictions show the use of Fry and Pagan's (2011) MT method. Section 5 concludes.

2. Identification Trough Sign Restrictions

To illustrate the idea behind recovering the structural shocks using sign restrictions, consider the following reduced-form VAR(1) model with n endogenous variables of the form

$$y_t = Ay_{t-1} + \varepsilon_t \text{ for } t = 1, 2, \dots, T \quad (1)$$

where y_t is an $n \times 1$ vector of variables, A is an $n \times n$ matrix of coefficients, and ε_t is a set of errors with mean zero, zero autocorrelation, and variance covariance matrix

$$\Sigma = E[\varepsilon_t \varepsilon_t'] \quad (2)$$

The reduced-form representation above summarises the sampling information in the data. ε_t is serially uncorrelated and orthogonal to the regressors in each equation. But ε_t has, in principle, no economic interpretation since the elements of ε_t still might be correlated across equations. Economic theory suggests that the one-step ahead forecast errors of a reduced-form VAR model are functions of

¹For an implementation of these methods see, for example, `vars` by Pfaff (2015).

some fundamental innovations, such that

$$B\varepsilon_t = e_t \quad (3)$$

where B is a $n \times n$ matrix of structural parameters and e_t are the structural shocks following a standard-Normal distribution with zero mean and a unit variance. Thus, the structural parameters can be recovered from

$$BB' = \Sigma = E[\varepsilon_t \varepsilon_t'] \quad (4)$$

$E[\varepsilon_t \varepsilon_t']$ can be obtained from estimating the reduced form VAR using Ordinary Least Squares (OLS). Recovering the structural shocks from \hat{e}_t requires identification of B . Since B contains n^2 unknown elements, identification of B requires at least $n(n-1)/2$ restrictions to uniquely identify the elements of B . The standard approach to this identification problem has been using a Cholesky decomposition or applying short-run or long-run restrictions to recover the structural shocks as suggested by e.g. Sims (1980), Blanchard and Quah (1989), and Gali (1992). These “hard” constraints are, for the most part, hard to reconcile with standard DSGE models. Instead of imposing hard restrictions on the model coefficients, sign restrictions only impose relatively weak prior beliefs on the responses of the form: “ x does not increase y for a certain period of time”.

While it is difficult to impose sign restrictions directly on the coefficient matrix of the model, it is easy to impose them ex-post on a set of orthogonalised impulse response functions. Thus, sign restrictions essentially explore the space of orthogonal decompositions of the shocks to see whether the responses conform with the imposed restrictions (Canova and De Nicolò, 2002).² In addition to making a choice about the signs of the responses, one has to specify for how long these restrictions apply after the impact of the shock. In theory, any length between only the first period after the shock and the entire length of response is possible.

The steps involved in recovering the structural shocks, given a set of sign restrictions, can be summarised as follows:

1. Run an unrestricted VAR in order to get \hat{A} and $\hat{\Sigma}$.
2. Extract the orthogonal innovations from the model using a Cholesky decomposition. The Cholesky decomposition here is just a way to orthogonalise shocks rather than an identification strategy.
3. Calculate the resulting impulse responses from Step 2.
4. Randomly draw an orthogonal *impulse vector* α
5. Multiply the responses from Step 3 times α and check if they match the imposed signs.
6. If yes, keep the response. If not, drop the draw.
7. Repeat Steps 2–6.

Step 1 is estimated using OLS. Step 1 as well as the calculation of the candidate impulse responses in Step 3 is based on modified versions of Sims’ (2010) **VAR tools**, which are considerably faster than other implementations currently available in R.

Steps 4 and 5 need some further explanation. The *impulse vector* essentially sets the loading of the shock onto the variables (Doan, 2011). Uhlig (2005) shows that a vector is an *impulse vector*, iff there is an n -dimensional vector a of unit length, i.e. $\|a\| = 1$ such that

$$\alpha = \tilde{B}a \quad (5)$$

where $\tilde{B}\tilde{B}' = \Sigma$ is a matrix decomposition of Σ . Uhlig (2005) shows that, given an impulse vector α , one can simply calculate the impulse responses by multiplying the impulse vector with the impulse responses obtained in Step 3.

The types of decompositions differ slightly across the three core procedures in **VARsignR**. In the case of the two procedures suggested by Uhlig’s (2005), the generation of the impulse vector is based on a Givens rotation, whereas in the case of RWZ (2010) it is based on a QR-decomposition. As a result, the draw for a to get the impulse vector α differs as well. In the case of Uhlig’s (2005) rejection method, one can directly draw directly an $n \times 1$ vector from the unit sphere, whereas in the case of the RWZ (2010) rejection method, the draw is an $n \times 1$ vector from the standard normal distribution. In the case of Uhlig’s (2005) penalty function method, a is based on a $(n-1) \times 1$ draw from a standard normal distribution and then projected into the R^n space using a stereographic projection. I will explain this in more detail below.

²Uhlig (2005) argues that, to some extent, sign restrictions make explicit restrictions which are often used implicitly by the researcher as in what she considers to be a reasonable response to a shock.

Sign restrictions are, for the most part, only well defined from a Bayesian point of view (Moon and Schorfheide, 2012). Steps 2–6 are based on a joint draw from a flat Normal Inverted-Wishart posterior for the VAR parameters and a uniform distribution for α as discussed above.³ One can then conduct inference based on the accepted draws and construct error bands similar to Sims and Zha (1999). While it might be desirable to have other priors than an uninformative one, Baumeister and Hamilton (2014) show that the algorithms for identifying sign restrictions only work for a very particular prior distribution, namely a uniform Haar prior and that standard informative priors will influence the posterior inference, even if the sample size is infinite. Thus, in the worst case scenario the researcher will simply analyse the prior distribution rather the posterior. A flat prior also has the advantage that standard (likelihood-based) specification tests, readily available in other packages, can be used and the researcher does not have to calculate, the marginal likelihood of the model. Lastly, a flat prior on the unit sphere is also appealing, since the results are unaffected by the Cholesky ordering in Step 2 (Uhlig, 2005).

The draws for the coefficients of the model from the multivariate Normal can be easily obtained using the `mvn` command of Fasiolo et al's (2015) `mvnfast` package. The Wishart draws are inverted draws from `rWishart` from the `stats` package. Draws from the unit sphere and the standard normal distribution to create α can be easily obtained using the `rballunif` command from the `HI` package by Petris and Tardella (2015) package and the `rnorm` function.

Step 5 involves constructing a suitable algorithm to check whether the impulse response functions have the appropriate sign or not. In the case of Uhlig's (2005) and RWZ's (2010) rejection methods, the algorithm consists of a number of sub-draws to generate α for each posterior draw. The algorithm then checks whether the imposed sign restrictions are satisfied for each restricted response and each restricted period, starting off with response of the shocked variable itself. In case, the response of the shocked variable and α in the first period both have the wrong sign, the signs of the impulses and α are flipped, and the algorithm checks whether or not the restrictions are met or not. If all restrictions for all restricted periods are met, the draw is kept and the function moves to the next posterior draw. Otherwise the draw is rejected. If the draw is rejected, the algorithm tries another sub-draw for α from the same posterior draw to find a combination of the impulses and α that matches the restrictions. If the algorithm cannot find a sub-draw for which the sign restrictions are met before reaching the maximum number of allowed sub-draws, it jumps to the next posterior draw after the maximum sub-draw.

In the case of Uhlig's (2005) penalty function, the algorithm is not based on the acceptance and rejection of sub-draws, rather than finding an impulse vector α which comes as close as possible to satisfying the imposed sign restrictions by minimising a function that penalises sign restriction violations. Let J be the total number of sign restrictions and K the total number response periods for which the restrictions apply. The impulse vector here is the vector α which minimises the total penalty $\Psi(\alpha)$ for all constrained responses $j \in J$ at all constrained response periods $k \in K$.

In order to treat all failures across the impulse responses symmetrically, the penalty function has to be adjusted for the scale of the variables and the sign of the restrictions. To treat the signs equally, let $l_j = -1$ if the sign of the restriction is positive and $l_j = 1$ if the restriction is negative. Scaling the variables is done by taking the standard error of the first differences σ_j of the variables as in Uhlig (2005).

Let $r_{j,\alpha}(k)$ be the response of j at response step k to the impulse vector α , then the minimisation problem can be written as

$$\min_{\alpha} \Psi(\alpha) = \sum_{j \in J} \sum_{k \in K} b \cdot f \left(l_j \frac{r_{j,\alpha}(k)}{\sigma_j} \right) \quad (6)$$

where b is a penalty depending on $f(\cdot)$ such that

$$b = \begin{cases} 1 & \text{if } f \left(l_j \frac{r_{j,\alpha}(k)}{\sigma_j} \right) \leq 0 \\ \text{penalty} & \text{if } f \left(l_j \frac{r_{j,\alpha}(k)}{\sigma_j} \right) > 0 \end{cases} \quad (7)$$

and *penalty* is a scalar ≥ 0 set by the user. While the specifications of $\Psi(\alpha)$ and b have no theoretical justification, Uhlig (2005) provides some discussion why the forms above are more suitable than others.

The next step is to find the impulse vector that minimises the total penalty for the restricted variables at all restricted horizons. Minimisation over the unit sphere of the above function is done as follows. First, I follow Doan (2011) and parametrise the unit sphere in n -space by randomly drawing $n - 1$ from a standard Normal-distribution and mapping the draw onto the n unit sphere using a

³Moon et al. (2011) apply sign restrictions using frequentist methods and show that the number of parameters to estimate can be very large and the topology of the identified set is quite complex.

stereographic projection. Second, I use the UOBYQA algorithm by [Powell \(2002\)](#) as implemented in [minqa](#) by [Bates et al. \(2015\)](#) to numerically find the minimum of the function.⁴

There are two important features of sign restrictions, the user should be aware of. First, all three procedures implemented in **VARsignR** only allow for partial identification of the model. Hence, only one shock can be fully identified. While the rejection methods by [Uhlig \(2005\)](#) and [Rubio-Ramírez et al. \(2010\)](#) can handle multiple shocks in theory, the penalty function approach, in its basic version, cannot. [Mountford and Uhlig \(2009\)](#) provide an extension of the penalty function approach that is able to cope with multiple shocks simultaneously.⁵ Second, the methods discussed here, especially the rejection methods, are part of a class of so-called set identified models ([Moon and Schorfheide, 2012](#)). This means that the parameters of interest are defined by a set of restrictions that do not point-identify them. Thus, the impulse responses of the models, at least in some cases, can only be bounded. I will discuss the implications of set-identification for inference and why partial identification is not necessarily a bad thing in Section 4.

3. Using VARsignR

3.1 Installation Instructions

VARsignR will be available on CRAN shortly. For now, I will demonstrate how to install the latest [GitHub](#) version of the package.

VARsignR depends on the [HI](#), [minqa](#), and [mvnfast](#) packages. In case you haven't installed the packages yet, run

```
install.packages("minqa")
install.packages("HI")
install.packages("mvnfast")
```

prior to installing **VARsignR**.

The simplest way to install **VARsignR** is to first download the latest (current version is 0.1.2) tarball [VARsignR_0.1.2.tar.gz](#) from [GitHub](#) and save it to your hard drive.

Run the following lines to install **VARsignR**

```
install.packages("PATH/VARsignR_0.1.2.tar.gz", repos = NULL)
```

where PATH has to be replaced with the location of [VARsignR_0.1.2.tar.gz](#) on your hard drive.⁶

3.2. Getting Started

To start off, type:

```
rm(list = ls())
set.seed(12345)
library(VARsignR)
data(uhligdata)
```

The first two lines clear the memory and set the seed for the random draws. These two lines are not essential, but will ensure that the user gets the exact same results as presented below. The third and the fourth line load the package and the example data set.

The examples below are based on the original data set used by [Uhlig \(2005\)](#). The data set contains monthly data for the United States (US) on Real GDP (y), GDP deflator (yd), commodity price index (p), FED funds rate (i), non-borrowed reserves (rnb), and total reserves (rt) from 1965:1–2003:12. The data have been taken from [Doan \(2011\)](#). Quarterly time series have been interpolated to monthly frequency using the interpolation methods by [Bernanke et al. \(1997\)](#) and [Bernanke and Mihov \(1998\)](#). For more details see `help(uhligdata)` and [Uhlig \(2005\)](#). All variables in the data set, except the FED funds rate, are the natural log of the original data series times 100 in line with the transformations of the data of the original paper.

⁴Note that the standard approach to minimising the penalty function is using a Nelder-Mead algorithm as implemented in [optim](#). Tests using both algorithms to minimise $\Psi(\alpha)$ showed that the UOBYQA algorithm has less convergence issues and is considerably faster without changing the results of the routine.

⁵This procedure is currently in development. Future releases of **VARsignR** will contain this procedure. Development files can be downloaded at <https://github.com/chrstdanne/VARsignR>.

⁶Under Windows, this will be something of the form "c:/docs/mydir". Under Mac and Linux, PATH will look like "/docs/mydir".

Uhlig's (2005) original question was to analyse the effect of an unanticipated monetary policy shock, i.e. an increase of one standard deviation in the FED funds rate on US real GDP and the so-called "price puzzle", i.e. an increase in domestic inflation in response to a contractionary interest rate shock which is at the odds with economic theory.⁷ Based on standard DSGE models, Uhlig (2005) suggested that an unanticipated innovation in the FED's policy rate

- does not decrease the FED's policy rate for x months after the shock
- does not increase commodity prices and inflation for x months after the shock
- does not increase non-borrowed reserves for x months after the shock

Thus, Uhlig (2005) uses four restrictions in total to identify a monetary policy shock. The 1st and the 6th variable of the model (real GDP and total reserves) remain unrestricted. Table 1 summarises the sign restrictions used to identify a monetary policy shock in the models below.

Shock/Variable	y	yd	p	i	rnb	rt
i	≤ 0	≤ 0	≤ 0	≥ 0	≤ 0	≤ 0

Table 1: Sign restrictions of the model

Given the ordering of the variables in the data set with y appearing first, followed by yd , p , i , rnb , and rt , one can specify these restrictions (on the 2nd, 3rd, 4th, and 5th variable) as a list of the form

```
constr <- c(+4,-3,-2,-5)
```

It is important to point out that the first element of `constr` indicates the shock of interest in the model. One *must* specify at least a sign restriction for the shock of interest. Any additional restrictions are optional and can appear in any order. The sign of the integers in `constr` indicate the direction of the restrictions, i.e. a "+" indicates a response ≥ 0 and "-" indicates a response restriction ≤ 0 .

3.3 Uhlig's (2005) Rejection Method

I start off by demonstrating Uhlig's (2005) rejection method. The model is estimated using the same settings as in Uhlig (2005), i.e. 12 lags, no constant, and 60 steps for the impulse response functions. I use 200 draws from the posterior and 200 subdraws for each posterior draw to generate the impulse vectors and the candidate impulse responses to which the rejection algorithm will be applied.

KMIN and KMAX denote the first and the last period of the responses to which the sign restrictions apply respectively. KMIN=1, KMAX=6 implies that the restrictions apply from the point of impact up to the 6th period of the response.⁸

nkeep denotes the desired number of accepted draws. This is different from the original implementation of Uhlig's (2005) routine. Uhlig (2005) used a large number of draws and sub-draws and used all the resulting accepted draws. Given that these routines can be quite time consuming, nkeep ensures that the routine stops, once a sufficient number of accepted draws is reached. In case the number desired draws is not reached before the maximum number of draws is reached, the routine stops when the maximum number of draws is reached and issues a warning.

To estimate the model, run

```
model1 <- uhlig.reject(Y=uhligdata, nlags=12, draws=200, subdraws=200, nkeep=1000, KMIN=1,
                      KMAX=6, constrained=constr, constant=FALSE, steps=60)
```

```
Starting MCMC, Tue Dec 8 09:36:53 2015.
```

```
|=====
```

```
| 44%
```

```
MCMC finished, Tue Dec 8 09:37:01 2015.
```

uhlig.reject returns an object containing the posterior draws of the coefficients of the model (BDraws), the variance-covariance matrix (SDraws), as well as posterior draws for the impulse response functions (IRFS), forecast error variance decompositions (FEVDS), and the implied shocks (SHOCKS) of the model. IRFS and FEVDS are arrays of dimensions $nkeep \times nstep \times nvar$. BDraws is $nkeep \times ncoef \times nvar$. SDraws is $nkeep \times nvar \times nvar$. SHOCKS is of $nkeep \times (T - nlags)$. In case, the estimation routine does

⁷The user may check the Appendix for a conventional VAR model identifying shocks using a Cholesky decomposition that produces the price puzzle.

⁸Note that any values satisfying $KMIN \leq KMAX \leq nstep$ will work. Also note that only restricting the first period of the responses, i.e. KMIN=1, KMAX=1 is computationally equivalent to the procedure by Faust (1998).

not produce the desired number of accepted draws, first dimension of the arrays will be equivalent to the number of accepted draws found by `uhlig.reject`. `summary(model1)` shows the output of `uhlig.reject`.

The number of rejected draws provides a crude indicator for how “well” the model is specified. A large number rejected draws, implies that there are potentially many other models that fit the data better than the one described by the current set of variables and/or sign restrictions.

```
summary(model1)
      Length Class  Mode
IRFS   360000 -none- numeric
FEVDS  360000 -none- numeric
SHOCKS 456000 -none- numeric
BDraws 432000 -none- numeric
SDraws  36000 -none- numeric
```

The standard way of analysing the results of a Bayesian VAR model is to take the median of the impulse response draws and plot them together with a pair user specified error bands or presenting the FEVDs of the model.

To do this, one can type:

```
irfs1 <- model1$IRFS

v1 <- c("GDP", "GDP Deflator", "Comm.Pr.Index", "Fed Funds Rate",
       "NB Reserves", "Total Reserves")

irfplot(irfdraws=irfs1, type="median", labels=v1, save=FALSE, bands=c(0.16, 0.84),
       grid=TRUE, bw=FALSE)
```

The first line above extracts the draws of the impulse response functions from the model results. The next line is not essential and only for convenience. `v1` defines a set of variable labels for the resulting graphs. In case no labels are specified, the variable names of the original data set will be used as plot titles. `irfplot` then plots the impulse responses using the standard 68 per cent error bands as specified in `bands`. `save=FALSE` indicates that the resulting graph will not be saved to the current working directory. `grid=TRUE` and `bw=FALSE` indicate that the plots will result grid lines and that coloured graphs instead of black and white graphs will be displayed. The user may want to check `help(irfplot)` for more information.

Fig. 1 shows the impulse response functions of the model.

In the same way, one can plot the FEVDs of the model in (per cent) for the shock of interest by typing⁹

```
fevd1 <- model1$FEVDS
fevdplot(fevd1, label=v1, save=FALSE, bands=c(0.16, 0.84), grid=TRUE,
       bw=FALSE, table=FALSE, periods=NULL)
```

Fig. 2 shows the resulting FEVDs of the model.

`fevdplot` has two extra options compared to `irfplot`, namely `table` and `periods`. In case, the user prefers to just show the FEVDs in a table, the `table=TRUE` option can be used. if `table=TRUE` and no value for `periods` is specified, `fevdplot` returns a table containing the FEVDs for all response steps. Specifying a list of values for `periods` with each entry \leq steps allows the user to generate a table of FEVDs for a selected number of periods. The example below, for example, generates a table containing the FEVDs for a monetary policy shock for every 10th response step.

```
fevd.table <- fevdplot(fevd1, table=TRUE, label=v1, periods=c(1,10,20,30,40,50,60))

print(fevd.table)
```

	GDP	GDP Deflator	Comm.Pr.Index	Fed Funds Rate	NB Reserves	Total Reserves
1	8.48	7.21	6.55	11.54	5.63	7.70
10	9.56	8.50	6.99	13.02	8.50	8.47
20	11.86	9.53	7.29	13.12	9.75	9.41
30	13.03	9.64	7.65	13.00	9.94	9.85
40	12.98	9.81	8.37	12.89	10.16	9.90

⁹Note that Uhlig (2005) does not show FEVDs, rather than forecast error revisions. Hence, the figure below looks different from Uhlig's (2005) analysis of variance.

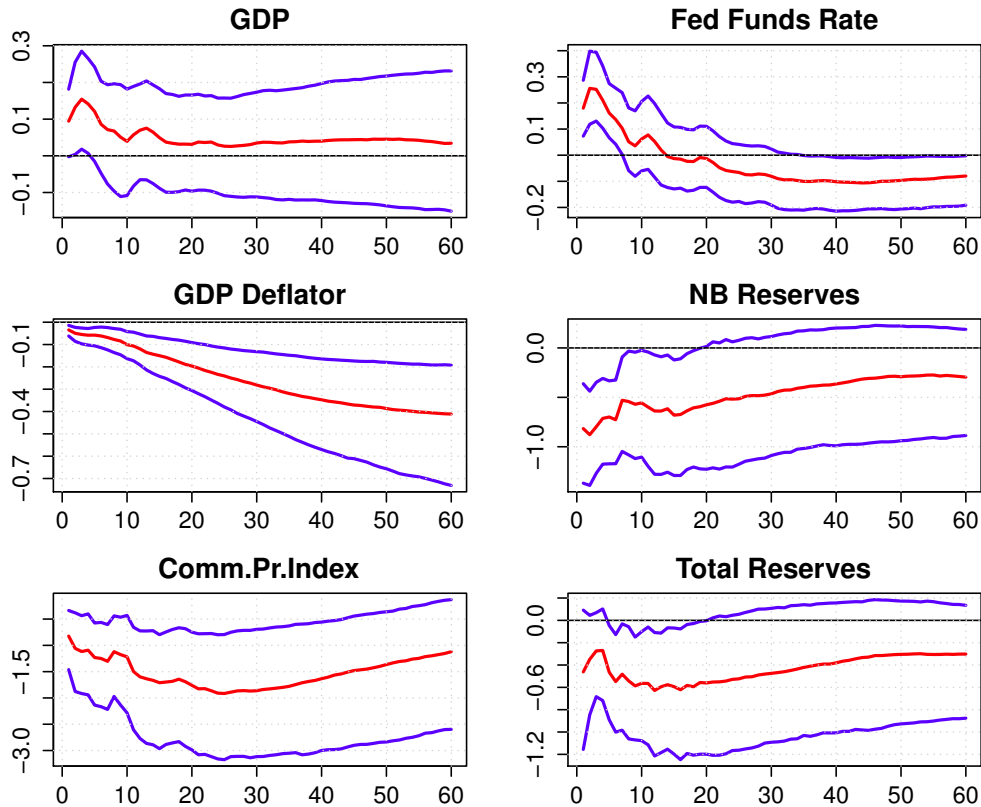


Figure 1: Replication of Fig. 6 in Uhlig (2005)

50	12.49	9.71	8.89	12.71	10.16	9.96
60	12.08	9.91	9.31	12.67	10.24	10.12

The user may also want to extract the series of shocks from the model by typing

```
shocks <- model1$SHOCKS
ss <- ts(t(apply(shocks,2,quantile,probs=c(0.5, 0.16, 0.84))), frequency=12, start=c(1966,1))

plot(ss[,1], type="l", col="blue", ylab="Interest rate shock", ylim=c(min(ss), max(ss)))
abline(h=0, col="black")
```

where the first line of the code block extracts the series of shocks, the second line calculates the median and the 68 per cent error bands of the shock, and the last two lines plot the median shock. Fig. 3 shows the implied series of shocks from the model.

The reader may use e.g.

```
lines(ss[,2], col="red")
lines(ss[,3], col="red")
```

to add the error bands to the plot below.

In the same fashion, one can also plot the posterior draws of the model (BDraws and SDraws), if desired (not shown here).

3.4 Rubio-Ramirez et al's (2010) Rejection Method

To illustrate RWZ's (2010) rejection method, the model is re-estimated using the same specifications and restrictions as above in Section 3.3. All options of `rwz.reject` are identical to the options of `uhlig.reject`.

To estimate the model using the RWZ (2010) approach, type

```
model2 <- rwz.reject(Y=uhligdata, nlags=12, draws=200, subdraws=200, nkeep=1000,
                    KMIN=1, KMAX=6, constrained=constr, constant=FALSE, steps=60)
```

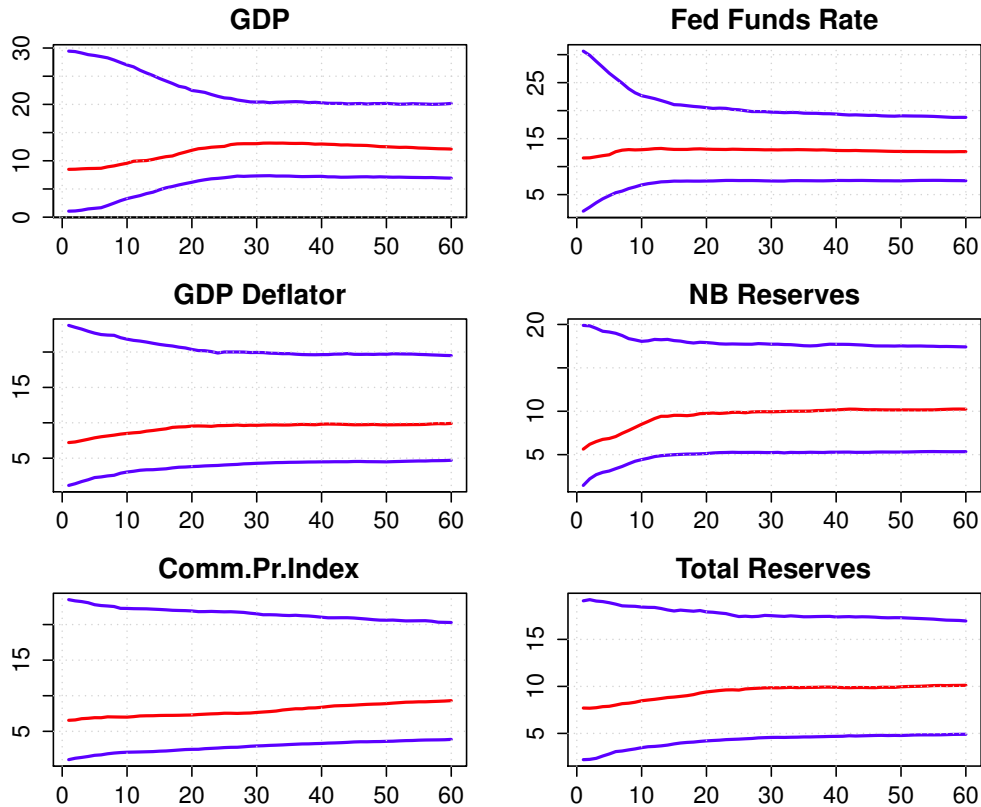


Figure 2: Variance decompositions for uhlig.reject

Starting MCMC, Tue Dec 8 10:08:31 2015.

|=====| 90%

MCMC finished, Tue Dec 8 10:08:49 2015.

```
irfs2 <- model2$IRFS
```

```
irfplot(irfdraws=irfs2, type="median", labels=v1, save=FALSE, bands=c(0.16, 0.84),
        grid=TRUE, bw=FALSE)
```

As shown by Fig. 4, the impulse response functions generated by `rwz.reject` are identical to the ones using `uhlig.reject` presented in the previous sub-section.

I skip the remaining results, since they are almost identical to the ones presented in the previous Section.

3.5 Uhlig's (2005) Penalty Function Method

One shortcoming of the two rejection methods above is that all impulse vectors satisfying the sign restrictions are considered to be equally likely (Uhlig, 2005). Moreover, by construction, the rejection methods will only find impulse vectors that will exactly satisfy the sign restrictions. In some cases, there might be only very few impulse vectors which satisfy a full set of restrictions. A way to tackle this problem is Uhlig's (2005) penalty function method. As discussed in Section 2, `uhlig.penalty` minimises a criterion function that penalises for sign violations.

The `uhlig.penalty` options `nlags`, `nstep`, `draws`, `KMIN`, `KMAX`, and `nstep` are the same as before. However, `subdraws` here refers to the maximum number of iterations of the UOBYQA algorithm. The `penalty` option specifies the penalty of $\Psi(\alpha)$ for sign violations as described in Section 2. Since, there are no sub-draws in the routine that can be used for inference, one needs a larger number of draws to generate the same number of desired acceptances.

To ensure near-certain convergence of minimisation problem, `uhlig.penalty` executes the minimisation algorithm twice for each draw of α . A draw here, gets rejected if either of the two minimisation

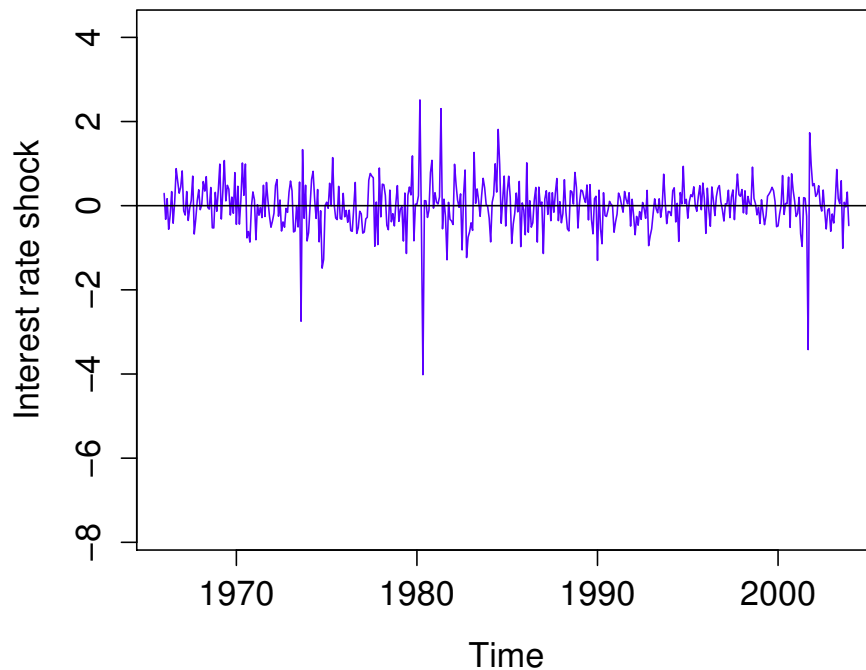


Figure 3: Interest rate shock

runs does not converge, or if the two runs reach different optima, i.e. different values of $\Psi(\alpha)$ for the same draw of α . crit here is the (absolute) threshold value for the difference between the first and the second run of the minimisation routine above which the draw gets rejected.

To estimate the model using Uhlig's (2005) penalty function, one can run

```
model3 <- uhlig.penalty(Y=uhligdata, nlags=12, draws=2000, subdraws=1000,
                        nkeep=1000, KMIN=1, KMAX=6, constrained=constr,
                        constant=FALSE, steps=60, penalty=100, crit=0.001)
```

```
Starting MCMC, Tue Dec 8 10:22:34 2015.
```

```
|=====
```

```
| 50%
```

```
Warning! 3 draw(s) did not converge.
```

```
MCMC finished, Tue Dec 8 10:24:49 2015.
```

In the above example, uhlig.penalty needs 1003 draws to generate 1000 accepted draws. 3 posterior draws were rejected since the UOBYQA algorithm did not converge in either of the two runs or converged to two different optima. Hence, for 0.3 per cent of the draws the algorithm could not find a minimum. This is an acceptable number and one can continue with the remainder of the analysis. In case there is a larger number of rejections due to non-convergence, the user may want to reconsider the model specification or the sign restrictions imposed on the responses.

The lines below extract the posterior impulse responses and plots the resulting impulse responses to a monetary policy shock.

```
irfs3 <- model3$IRFS
irfplot(irfdraws=irfs3, type="median", labels=v1, save=FALSE, bands=c(0.16, 0.84),
        grid=TRUE, bw=FALSE)
```

Fig. 5 shows the resulting graph.¹⁰

¹⁰Note that Uhlig (2005) only uses 100 accepted draws (e.g. nkeep=100, draws=200) to construct the impulse responses of the penalty function in his paper. Due to the larger number of draws, the responses look slightly smoother than the ones shown in Uhlig (2005).

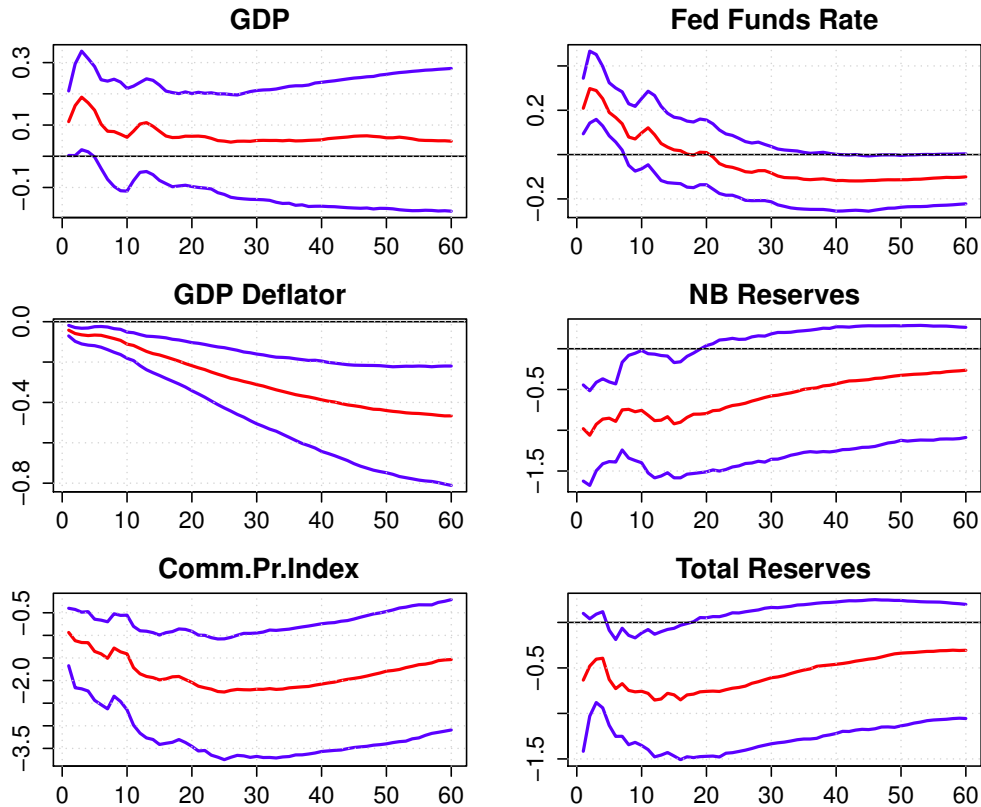


Figure 4: Replication of Fig. 6 in Uhlig (2005) using Rubio-Ramirez et al's (2010) method

4. Critique and Issues

While sign restrictions provide a very appealing way of recovering the structural shocks from an economic point of view, several papers have pointed out a number of shortcomings of sign restrictions. Addressing these issues is subject to a lot of current research in macroeconometrics.

4.1 Multiple Shocks

VARsignR currently allows only partial identification of the model. Partial identification is not necessarily a bad thing. The key idea behind sign restrictions is to characterise a shock through placing restrictions on the responses of some variables, but being agnostic about others. While it is possible to identify all shocks of the model, doing so by just using sign restrictions is inherently difficult. As shown by [Fry and Pagan \(2011\)](#) there are many cases in the literature, where researchers aim to identify all shocks in the model but fall short of doing so. One reason for this, is that different shocks in the model might be characterised by the same set of restrictions. Considering the data set at hand, it is not too hard to imagine a model in which a GDP deflator shock is characterised by the same restrictions as a shock to the commodity price index. Another reason for underidentification is that researchers do not provide enough restrictions to uniquely identify all shocks in the model.

Thus, focussing only on one shock of the model and being explicit about partial identification might be a better way of approaching a particular research question.

4.2 Set Identification and Inference

As pointed out in Section 2, models identified by sign restrictions are only set-identified. This means that, while sign restrictions solve the structural identification more elegantly than e.g. a recursive structure, sign restrictions might not necessarily generate a unique set of impulse responses. Depending on the problem at hand, sign restrictions may generate a new set of structural equations and shocks for each rotation of α which means each draw produces a set of possible inferences, each of which is equally consistent with both the observed data and the underlying restrictions ([Moon et al., 2011](#)). Hence, summarising the responses using e.g. the median response and conventional error bands to represent the spread of responses displays the distribution across models and not, as

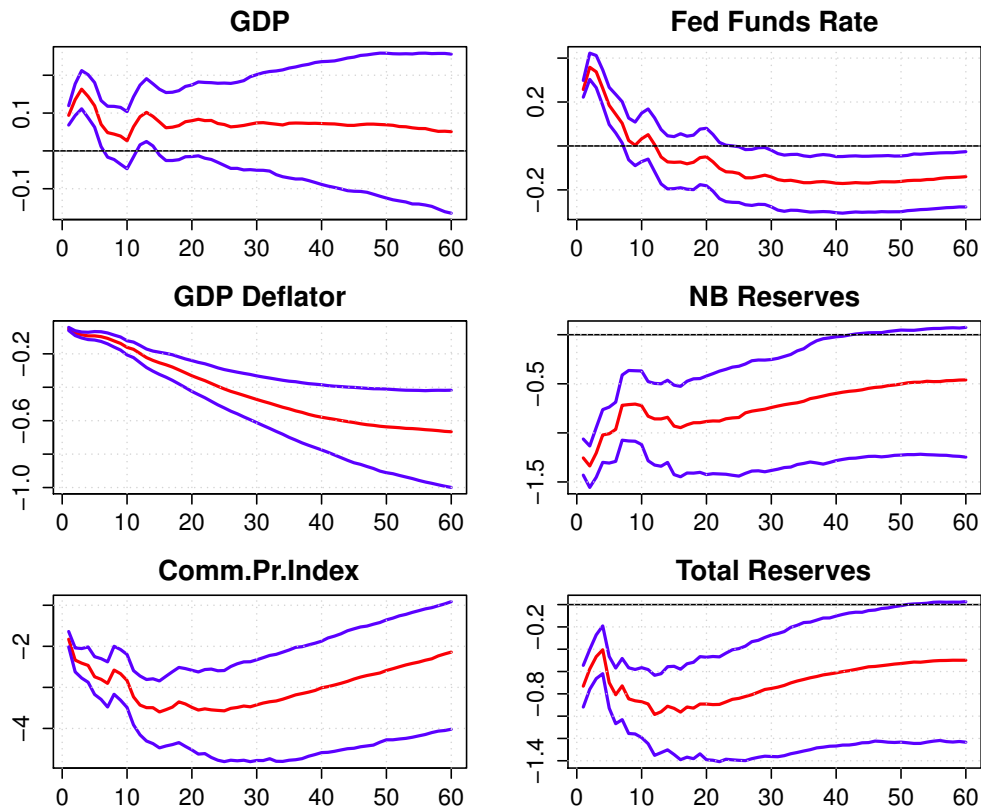


Figure 5: Replication of Fig. 14 in Uhlig (2005)

intended, sampling uncertainty. Uhlig's (2005) and RWZ's (2010) rejection method are particularly prone to this "model identification" problem. While the question of whether there is *any* model for which the data can produce certain responses to shocks is an interesting one in itself, it is unsatisfactory in most cases.

There are several ways to deal with this problem. One way is to narrow down the set of admissible models to a singleton (Fry and Pagan, 2011). This is what Uhlig's (2005) penalty function does by generating a "weighted sample" of all draws, including the ones that do not meet the imposed restrictions. Thus, if the results differ markedly across the three routines, the user should probably opt for `uhlig.penalty`.

A way to improve the results of `uhlig.reject` and `rwz.reject` is to use additional information and additional constraints. In general, the more restrictions one imposes, the "better" identified the model is, provided that the additional restrictions make sense (Paustian (2007)). Uhlig (2005) and Arias et al. (2014) discuss ways of how to strengthen sign restrictions by imposing additional zero restrictions on selected impact responses. Imposing zero restrictions, however, brings back the problems discussed in the introduction which Faust (1998), Canova and De Nicolo (2002), and Uhlig (2005) set out to solve.

To demonstrate how additional restrictions can improve the model's results, consider the example shown in Section 3.3 with four sign restrictions. As shown below, imposing an additional restriction on the response of total reserves (6th variable in the model) to a monetary policy shock (see `constr5` below) yields impulse responses that are closer to the ones produced by `uhlig.penalty` (see Fig. 6).

```
constr5 <- c(+4,-3,-2,-5,-6)

model5c <- uhlig.reject(Y=uhligdata, nlags=12, draws=200, subdraws=200, nkeep=1000,
                        KMIN=1, KMAX=6, constrained=constr5, constant=FALSE, steps=60)

Starting MCMC, Fri Dec 11 22:59:28 2015.
|=====| 62%

MCMC finished, Fri Dec 11 22:59:37 2015.

irf5c <- model5c$IRFS
irfplot(irf5c, labels=v1)
```

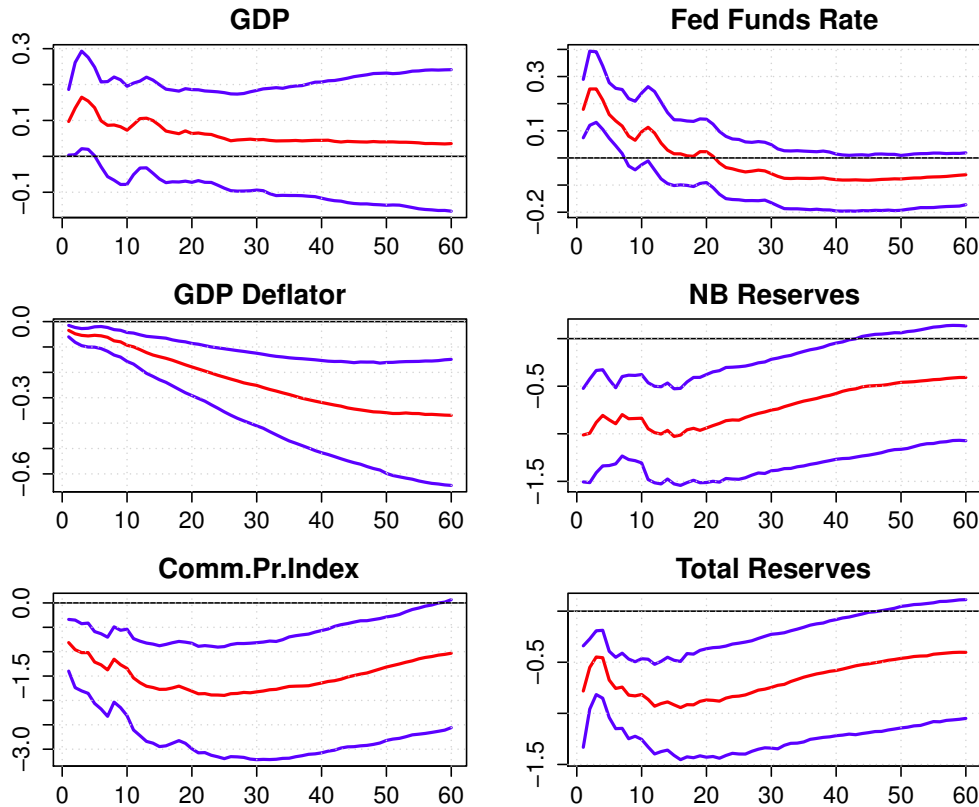


Figure 6: `uhlig.reject` with additional sign restrictions

In the same fashion, providing restrictions for the full set of shocks can solve the model identification problem as well. Canova and Paustian (2011) use simulated data from a DSGE model and show that sign restrictions recover the shocks of the model reasonably well, provided that enough restrictions are used and that all shocks are identified. This however, might not be feasible in many cases as discussed in Section 4.1. Kilian and Murphy (2012) suggest an alternative way of achieving proper identification without imposing zero restrictions or identifying the full set of shocks by setting upper bounds for certain parameters of the model.

One way to analyse how significant the model identification problem is and how well the shocks in the model are identified, is Fry and Pagan's (2011) Median-Target (MT) method. Fry and Pagan's (2011) MT method is to find the *single* impulse vector α that produces impulse responses that are as close to the median responses as possible. The MT method is essentially a diagnostic device. Fry and Pagan (2011) show that strong differences between the MT impulse responses and the median responses indicate the standard model inference will be biased and misleading.

The MT method finds the single best draw for α by minimising the sum of squared standardised gaps between the impulse responses given the test rotation and the sign restricted responses of the model that is tested. Minimisation of the target function is, as in `uhlig.penalty` based on the UOBYQA algorithm by Powell (2002).

Fry and Pagan's (2011) MT method is implemented in `fp.target`. The lines below demonstrate the usage of `fp.target` using the responses from the demonstration of `uhlig.reject` from Section 3.3 stored in `irfs1`. To run the MT method, type

```
fp.target(Y=uhligdata, irfdraws=irfs1, nlags=12, constant=F, labels=v1, target=TRUE,
         type="median", bands=c(0.16, 0.84), save=FALSE, grid=TRUE, bw=FALSE,
         legend=TRUE, maxit=1000)
```

The `irfdraws`, `nlags`, `constant`, and `labels` options have the same definitions as in the routines discussed above. The same applies to `bands`, `save`, `grid`, and `bw`. `type` is, similar to the `irfplot` option, whether the target for the MT method should be calculated with respect to the mean or the median of the responses. This, together with `nlags` and `constant`, should correspond to what the user has used before to generate the impulse responses, handed over to `irfdraws`, for consistency reasons. `target=TRUE` indicates that the target response, i.e. the response of the original model should appear in the resulting plots or not. `legend=TRUE` indicates that a legend is displayed at the bottom of the graph. `maxit` is the maximum number of iterations of the UOBYQA algorithm. In case the algorithm

does not converge, a warning is issued. The user may want to increase the number iterations in this case, or reconsider the original model specification.

Fig. 7 shows the resulting graph of `fp.target`. Since `target=TRUE`, the graph essentially compares the responses of `uhlig.reject` and the results of `fp.target`. It's important to point out that the impulse responses of `fp.target` do not have to lie within the error bands of the initial model, as for example it is the case for the response of NB reserves to a monetary policy shock. The gap here, suggests that median impulse response cannot be seen as a response of a single model and an alternative model should be considered. On the other hand, the remaining MT responses match the responses of `uhlig.reject` quite well in the long run. Something that one might consider evidence in favour of the current model specification.

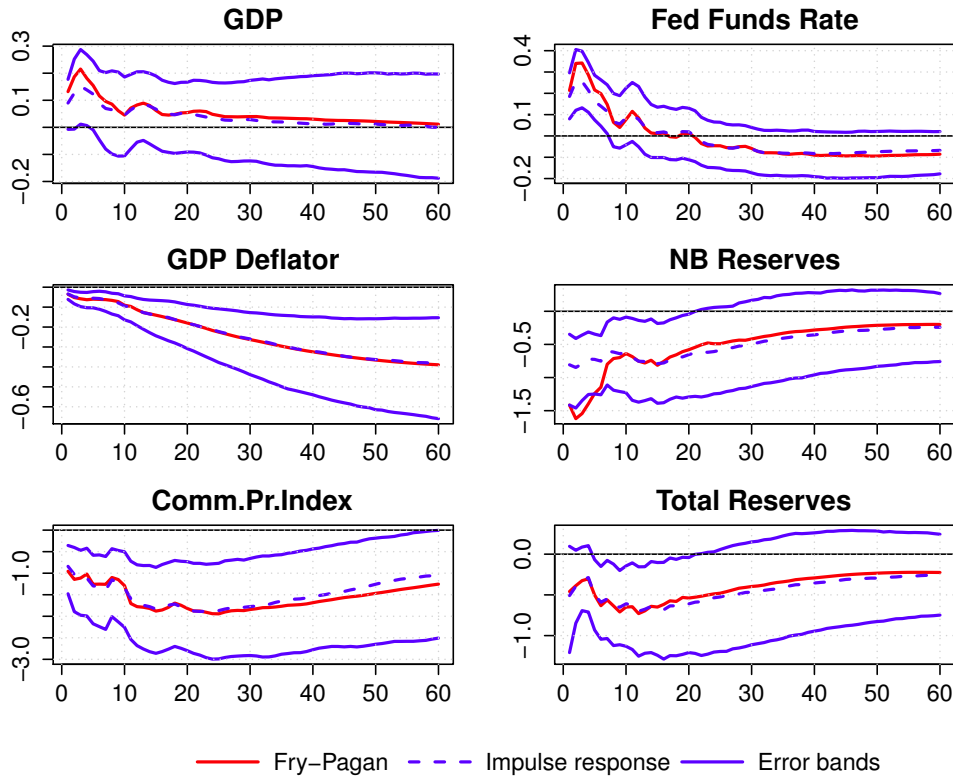


Figure 7: Fry and Pagan's (2011) median target method for `uhlig.reject`

5. Summary

This paper introduces the **VARsignR** package, an R package for estimating VAR models using sign restrictions. **VARsignR** implements the rejection method and penalty function method suggested by Uhlig (2005) and the rejection method developed by Rubio-Ramirez et al. (2010). The paper discusses the details of the implementation of the algorithms and demonstrates the usage of the routines. The paper also discusses some shortcomings of sign restrictions the researcher should be aware of when using **VARsignR** and demonstrates the use of Fry and Pagan's (2011) median target method as a diagnostic device.

Acknowledgements

VARsignR is, to a large extent, an adaptation of Tom Doan's (2011) RATS codes and Tao Zha's (2010) Matlab codes. In addition, it takes advantage of Chris Sims' (2010) VAR tools. I would like to thank them for making their programme codes available online.

Bibliography

- Arias, J., Rubio-Ramirez, J., and Waggoner, D. (2014). Inference based on svars identified with sign and zero restrictions: Theory and applications. *Federal Reserve System International Finance Discussion Papers* 1100. [p11]
- Bates, D., Mullen, K., Nash, J., and Varadhan, R. (2015). *minqa: Derivative-free optimization algorithms by quadratic approximation*. URL: <https://cran.r-project.org/web/packages/minqa/>. [p4]
- Baumeister, C. and Hamilton, J. (2014). Sign restrictions, structural vector autoregressions, and useful prior information. *NBER Working Paper No. 20741*. [p3]
- Bernanke, B., Gertler, M., and Watson, M. (1997). Systematic monetary policy and the effects of oil price shocks. *Brookings Papers on Economic Activity*, 1:91–157. [p4]
- Bernanke, B. and Mihov, I. (1998). Measuring monetary policy. *Quarterly Journal of Economics*, 113:869–902. [p4]
- Blanchard, O. and Quah, D. (1989). The dynamic effects of aggregate demand and supply disturbances. *American Economic Review*, 79:655–673. [p1, 2]
- Canova, F. (2007). *Methods for applied macroeconomic research*. Princeton University Press., Princeton, NJ. [p1]
- Canova, F. and De Nicolo, G. (2002). Monetary disturbances matter for business fluctuations in the G-7. *Journal of Monetary Economics*, 49:1131–1159. [p1, 2, 11]
- Canova, F. and Paustian, M. (2011). Business cycle measurement with some theory. *Journal of Monetary Economics*, 58:345–361. [p12]
- Canova, F. and Pina, J. (2005). What VAR tells about DSGE models? In Deibolt, C. and Krytsou, C., editors, *New Trends and Macroeconomics*. Springer, New York, NY. [p1]
- Del Negro, M. and Schorfheide, F. (2011). Bayesian macroeconometrics. In Geweke, J., K. G. and Van Dijk, H., editors, *The Oxford Handbook of Bayesian Econometrics*, page 293–389. Oxford University Press, Oxford: UK. [p1]
- Doan, T. (2011). *RATS programs to replicate Uhlig’s VAR identification technique*. URL: https://www.estima.com/procs_perl/uhligjme2005.zip. [p2, 3, 4]
- Fasiolo, M., van den Berg, T., Maurer, J., and Watanabe, S. (2015). *mvnfast package*. URL: <https://cran.r-project.org/web/packages/mvnfast>. [p3]
- Faust, J. (1998). On the robustness of the identified VAR conclusions about money. *Carnegie-Rochester Conference Series on Public Policy*, 49:207–244. [p1, 11]
- Fry, R. and Pagan, A. (2011). Sign restrictions in structural vector autoregressions: A critical review. *Journal of Economic Literature*, 49:938–960. [p1, 10, 11, 12, 13]
- Gali, J. (1992). How well does the IS-LM model fit postwar U.S. data. *Quarterly Journal of Economics*, 107:709–738. [p1, 2]
- Kilian, L. and Murphy, D. (2012). Why agonistic sign restrictions are not enough: Understanding the dynamics of oil market VAR models. *Journal of the European Economic Association*, 10:1166–1188. [p12]
- Moon, H. and Schorfheide, F. (2012). Bayesian and frequentist inference in partially identified models. *Econometrica*, 80:755–782. [p3, 4]
- Moon, H., Schorfheide, F., and Granziera, E. (2011). Inference for VARs identified with sign restrictions. *NBER Working Paper No. 17140*. [p3, 10]
- Mountford, A. and Uhlig, H. (2009). What are the effects of fiscal policy shocks? *Journal of Applied Econometrics*, 24:960–992. [p4]
- Paustian, M. (2007). Assessing sign restrictions. *B.E. Journals of Macroeconomics*, 7:1–33. [p11]
- Petris, G. and Tardella, L. (2015). *HI: Simulation from distributions supported by nested hyperplanes*. URL: <https://cran.r-project.org/web/packages/HI/>. [p3]
- Pfaff, B. (2015). *vars: VAR modelling*. URL: <https://cran.r-project.org/web/packages/vars/>. [p1]

- Powell, M. (2002). UOBYQA: unconstrained optimization by quadratic approximation. *Mathematical Programming - Series B*, 92:555–582. [p4, 12]
- Rubio-Ramirez, J., Waggoner, D., and Zha, T. (2010). Structural vector autoregressions: Theory of identification and algorithms for inference. *Review of Economic Studies*, 77:665–696. [p1, 2, 3, 4, 7, 11, 13]
- Sims, C. (1980). Macroeconomics and reality. *Econometrica*, 48(1-48). [p1, 2]
- Sims, C. (2010). *VAR tools*. URL: <http://sims.princeton.edu/yftp/VARtools/>. [p2]
- Sims, C. and Zha, T. (1999). Error bands for impulse responses. *Econometrica*, 67:1113–1156. [p3]
- Uhlig, H. (2005). What are the effects of monetary policy on output? results from an agnostic identification procedure. *Journal of Monetary Economics*, 52:381–419. [p1, 2, 3, 4, 5, 6, 8, 9, 11, 13, 15]

Author

Christian Danne
Dublin
Ireland
dannec@tcd.ie

Appendix

VARsignR also contains a function that allows the user to identify shocks using a standard Cholesky decomposition. The model is, as above, a Bayesian VAR model with a flat Normal Inverted-Wishart prior. The ordering of the Cholesky decomposition corresponds to the ordering of the variables in the data set. The user has to specify the shock of interest in shock. The number corresponds to the position of the variable in the underlying data.frame. To estimate a model comparable to the ones above, simply type

```
model0 <- rfbvar(Y=uhligdata, nlags=12, draws=1000, constant=FALSE, steps=60, shock=4)

irfs0 <- model0$IRFS

irfplot(irfdraws=irfs0, type="median", labels=v1, save=FALSE, bands=c(0.16, 0.84),
        grid=TRUE, bw=FALSE)
```

Note that the specification corresponds to the model presented in Figure 5 of Uhlig (2005). The object returned by `rfbvar` has the same structure as the models discussed in Section 3. Fig. A.1 below shows the resulting impulse response functions. As shown by the graph, the results differ markedly from the sign restriction procedure. The positive response of the GDP deflator to a contractionary monetary policy shock is known as the “price puzzle” which is at the odds with standard DSGE models.

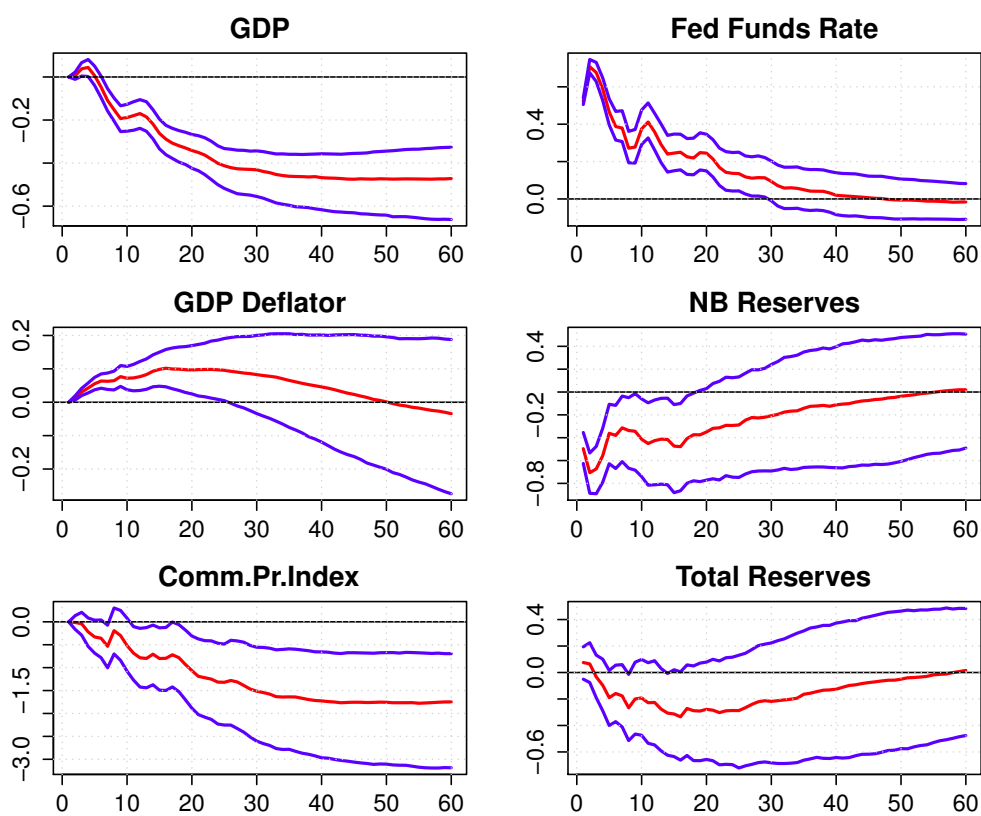


Figure A.1: Replication of Fig. 5 in Uhlig (2005)