Stargazing with Structural VARs: Shock Identification via Independent Component Analysis

DMITRY KULIKOV
and
ALEKSEI NETŠUNAJEV

Eesti Pank
Estonia pst. 13
15095 Tallinn
Estonia
phone: +372 6680777
e-mail: dmitry.kulikov@eestipank.ee

This version:
March 17, 2017

Abstract

A new ICA–based statistical identification procedure is introduced for Bayesian SVAR models with independent non–Gaussian structural innovations. Additional statistical information, available in non–normally distributed shocks, allows us to estimate a fully general structural VAR model without use of any strong conventional *a priori* identifying restrictions. The new procedure is validated using the US macroeconomic data series, where the nature of four different shocks is empirically examined. In particular, we statistically identify the short-run impacts and impulse responses of four structural shocks, which we label a "monetary policy shock", a "money demand shock", an "aggregate demand shock" and an "aggregate supply shock". We find a robust and well-pronounced "price puzzle" in response to our "monetary policy shock", while the statistically identified "money demand shock" induces a strong reaction in the US real output and prices. In addition, we also document a new "real output puzzle" in response to the Federal Reserve monetary policy action.

JEL CLASSIFICATION: C11, C32, C54
KEYWORDS: Independent component analysis, Non–Gaussian innovations, Statistical identification, Bayesian inference, Clustering methods, SVAR analysis

1 Introduction

Structural vector autoregressions (SVARs) are back in vogue after the global financial crisis as a tool for agnostic macroeconomic data exploration, unencumbered with the theoretical orthodoxy of fully–specified structural macroeconomic models. However, conventional approaches to the identification of SVAR models still rely on strong *a priori* restrictions on the system matrices, among which the short–run zero restrictions on the contemporaneous effects matrix (the structural matrix) remain the most widely used in the applied literature. Some of the most contentious critiques of the applied SVAR literature focuses on the nature and suitability of these *a priori* identifying restrictions.

Recently, a growing number of econometric studies has focused on the statistical identification of SVAR models, where the maximum effort is expended to relax all strong conventional *a priori* identification restrictions on the system matrices, including the structural matrix, and let the data decide freely on the conditional mean parameters of the model and the shape of impulse response functions. These studies, exemplified by Lanne and Lütkepohl (2008), Lanne, Lütkepohl and Maciejowska (2010), Lütkepohl and Netšunajev (2014), Herwartz and Lütkepohl (2014), and Kulikov and Netšunajev (2015) among others, effectively shift the burden of structural identification from the restrictions on system matrices to certain assumptions on the nature of fundamental structural innovations in SVAR models. The most popular of these statistical identification approaches had so far relied on the time–varying volatility of the structural shocks, along the lines of Rigobon (2003). This statistical identification approach presupposes the existence of coincident volatility shifts in the long spans of macroeconomic data in order to be able to successfully identify the structural innovations and their dynamic effects from the data. While in many circumstances the time–varying volatility assumption is tenable and had clearly been established on the empirical grounds, there remains a potentially large set of macroeconomic data series where the volatility shifts might be hard to detect or can be ruled out. This especially concerns macroeconomic datasets which cover relatively short time spans and/or tranquil historical periods.

One possible way of relaxing the time–varying volatility assumption on the structural innovations in SVAR models, while still being able to identify the system without strong *a priori* restrictions on the system matrices, is to abandon the widely used normality assumption on the distribution of structural innovations, at the same time keeping the crucial premise of independent policy innovations in the system. It turns out, that independent non–Gaussian innovations contain sufficient statistical information to be able to identify the system without the help of conventional identification restrictions and without an explicit assumption about the coincident volatility shifts, although the latter case is subsumed by the new framework.

In the technical literature, these ideas had been in development since the early 1980s under the banners of "blind source separation" and "independent component analysis". The problem of separating independent signals from their observed linear mixtures without the explicit knowledge of the mixing matrix had arisen in 1980s in the context of neural modelling and telecommunications. By the early 1990s the theoretical solution of this problem has been found and first working algorithms started to appear. By the early 2000s the techniques of blind source separation and independent component analysis had found a large number of applications outside their original fields and are currently undergoing a very active development in the technical literature. A comprehensive overview of this field can be found in Comon and Jutten (2010).

In this paper we explore the new ideas provided by blind source separation techniques for identification of Bayesian SVARs based on the criteria of independence and non–Gaussianity of the fundamental structural innovations. This criterium also relaxes the widely used normality assumption on the distribution of structural innovations. The idea is based on the independent component analysis (ICA) stemming from a spectrum of applications in the telecommunications, image analysis, neural networks and other technical fields. The first well–known application of these ideas to identification in SVAR models appeared in Hyvärinen, Zhang, Shimizu and Hoyer (2010). In the econometric literature, our paper is closely related to the ideas in Moneta, Entner, Hoyer and Coad (2013), Herwartz (2016) and Lanne, Meitz and Saikkonen (2017). Our paper differs from these previous contributions in two dimensions. Firstly, we remain faithful to the Bayesian statistical tradition and show how to achieve a fast and reliable posterior sampling with the ICA–based shock identification for SVAR models. Secondly, we fully espouse the concept of statistical shock identification, whereby the economic interpretation of statistically identified shocks comes from their estimated effects on the observed macroeconomic dynamics, rather than being imposed *a priori* as a part of the shock identification procedure.

The paper is organised as follows: Section 2 gives an overview of the structural model and the list of assumptions necessary for its identification using the ICA–based approach. Section 3 provides an overview of the Bayesian inference for SVAR models with independent non–Gaussian innovations, including the novel clustering–based approach for disentangling the order of structural innovations from the posterior parameter draws. Section 4 illustrates the new identification approach using a sample of the US macroeconomic data series, where four distinct structural shocks are uncovered by the ICA–based technique without any strong conventional *a priori* identifying restrictions on the system matrices.

2 SVAR model with independent non–Gaussian innovations

In this section we introduce the general class of SVAR models with independent innovations and provide an overview of the identification strategy for these models. In particular, we show how the identification can be achieved without the help of strong conventional *a priori* identification restrictions on the system matrices using the ideas from the ICA literature.

In order to describe the observed dynamics of macroeconomic variables in a relatively agnostic way, i.e. without imposing too much *a priori* structure on their evolution, we rely on the ideas from identified VAR literature, originating from the groundbreaking study of Sims (1980). We assume that the time evolution of an $n \times 1$ vector $\boldsymbol{y}_t$ of endogenous variables can be adequately represented by the following structural (SVAR) model:

$$\mathbf{A}_0 \boldsymbol{y}_t = \boldsymbol{k}_0 + \boldsymbol{k}_1 t + \mathbf{A}_1 \boldsymbol{y}_{t-1} + \ldots + \mathbf{A}_p \boldsymbol{y}_{t-p} + \boldsymbol{\epsilon}_t \,, \tag{1}$$

where $\boldsymbol{k}_0, \boldsymbol{k}_1$ are the optional vectors of deterministic parameters, $\mathbf{A}_0$ is the contemporaneous impact (structural) matrix, $\mathbf{A}_1, \ldots, \mathbf{A}_p$ are the autoregressive (convolution) matrices, and $\boldsymbol{\epsilon}_t$ is the vector of independent non–Gaussian structural innovations. The components of $\boldsymbol{\epsilon}_t$ may not all have identical marginal distributions, but each is required to be symmetric around zero and scale normalised to one. Collectively, $\mathbf{A}_0, \ldots, \mathbf{A}_p$ are referred to as the system matrices of the model. In the SVAR model, the innovations $\boldsymbol{\epsilon}_t$ are usually interpreted as "fundamental" through a prism of certain macroeconomic theories, and it is their effect on the dynamics of endogenous variables in $\boldsymbol{y}_t$ that is of a primary interest to a researcher using these models for applications in empirical macroeconomics and for the development of new theories; refer to Leeper, Sims and Zha (1996) for a general exposition.

In many ways, the structural matrix $\mathbf{A}_0$ is of central importance in this model. It determines the contemporaneous interactions of variables in $\boldsymbol{y}_t$ and plays a crucial role in shaping the impulse responses of the endogenous variables to the fundamental innovations. However, under the usual normality assumption on the structural innovations $\boldsymbol{\epsilon}_t$, the full identification of all system matrices $\mathbf{A}_0, \ldots, \mathbf{A}_p$ is not statistically feasible without some strong *a priori* restrictions on their elements. A textbook treatment of the conventional SVAR models with Gaussian innovations and the associate identification issues can be found in Hamilton (1994), Amisano and Giannini (1997) and Lütkepohl (2005).

In this paper we break away from the usual Gaussianity hypothesis on the distribution of structural innovations $\boldsymbol{\epsilon}_t$ and instead allow for non–normality of the structural innovations. Crucially, we retain the foundational assumption of independent structural innovations in $\boldsymbol{\epsilon}_t$, since any economic interpretation of the structural model (1) without this assumption would be rather different from the framework of Sims (1980) and all subsequent studies in the identified VAR literature. On the other hand, normality of the fundamental innovations $\boldsymbol{\epsilon}_t$ does not usually follow from any structural macroeconomic theories, bus is rather assumed on the statistical grounds. In this paper, we depart from the usual normality assumption because non–Gaussian independent shocks carry additional statistical information which helps to identify the structural model without resorting to the strong conventional *a priori* identification restrictions on the system matrices $\mathbf{A}_0, \ldots, \mathbf{A}_p$.[1]

Assuming that the structural matrix $\mathbf{A}_0$ is non-singular, the model can be written in the reduced (VAR) form, suitable for the likelihood–based statistical inference on the system matrices $\mathbf{A}_0, \ldots, \mathbf{A}_p$ and other model parameters:

$$\boldsymbol{y}_t = \boldsymbol{c}_0 + \boldsymbol{c}_1 t + \boldsymbol{\Phi}_1 \boldsymbol{y}_{t-1} + \ldots + \boldsymbol{\Phi}_p \boldsymbol{y}_{t-p} + \boldsymbol{u}_t \,, \tag{2}$$

[1]Notice that the conventional SVAR models with Gaussian innovations always assume independence of the structural innovations by postulating the variance–covariance matrix of $\boldsymbol{\epsilon}_t$ to be equal to unity. Hence, the new assumption of independent non–Gaussian shocks in (1) can be viewed as a natural extension of the accepted modelling framework.

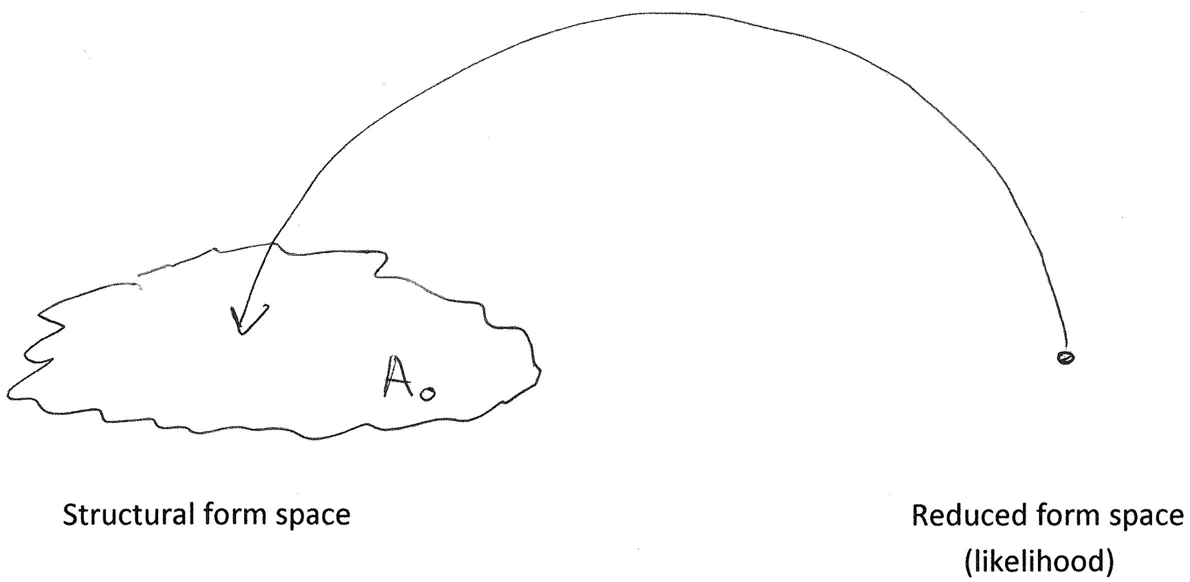

Figure 1: Identification in structural VAR models

where $\boldsymbol{c}_i = \mathbf{A}_0^{-1}\boldsymbol{k}_i$ for each $0 \leq i \leq 1$, $\boldsymbol{\Phi}_j = \mathbf{A}_0^{-1}\mathbf{A}_j$ for each $1 \leq j \leq p$, and where the vector of reduced form residuals is given by:

$$\boldsymbol{u}_t = \mathbf{A}_0^{-1}\boldsymbol{\epsilon}_t \, . \tag{3}$$

The reduced form deterministic parameters $\boldsymbol{c}_0, \boldsymbol{c}_1$, the autoregressive matrices $\boldsymbol{\Phi}_1, \ldots, \boldsymbol{\Phi}_p$ and the variance–covariance matrix of $\boldsymbol{u}_t$ can be directly inferred using Bayesian methods of combining priors and data (likelihood). In this paper we then proceed to identify the structural parameters $\mathbf{A}_0, \ldots, \mathbf{A}_p$ based on the reduced form statistical evidence and assumptions about the fundamental innovations $\boldsymbol{\epsilon}_t$, as shown diagrammatically in Figure 1.[2]

The reduced form residuals (3) in the reduced form model can be viewed as a linear mixture of independent structural innovations $\boldsymbol{\epsilon}_t$. The familiar identification problem manifests itself in the loss of statistical information related to the mixing matrix $\mathbf{A}_0^{-1}$ when the structural innovations $\boldsymbol{\epsilon}_t$ are normally distributed — only $\frac{n(n+1)}{2}$ elements of the structural matrix $\mathbf{A}_0$ can be estimated in this case. This result dates back to Darmois (1953) and is largely motivating the widespread use of the strong conventional *a priori* identification restrictions in the econometric literature, for without such restrictions the structural model would remain unidentified from the reduced form (likelihood) perspective given by (2). The familiar SVAR identification problem in the case of normally distributed fundamental innovations $\boldsymbol{\epsilon}_t$ can be succinctly rep-

[2]Alternatively, one can proceed in the opposite direction: from the space of structural models to the space of reduced form models. The comprehensive treatment of this approach can be found in Rubio-Ramírez, Waggoner and Zha (2010) for the case of strong conventional *a priori* identification restrictions. However, when the statistical identification of SVAR is desired, as is the cased in the current paper, the full exploration of the structural space becomes quite complicated, given the essential multiplicities of the observationally equivalent structural models and the lack of strong priors on $\mathbf{A}_0$. A recent study by Lanne et al. (2017) offer a possible perspective on the this issue.

resented by the one–to–many correspondence between the reduced form model space, for which the data evidence is summarised by the likelihood function, and the uncountable subspace of observationally equivalent structural models (1), each one with a different structural matrix $\mathbf{A}_0$, as shown in Figure 1.

All existing identification approaches in SVAR literature can therefore be viewed as a choice problem in the structural space — equivalently as a selection of a particular structural matrix $\mathbf{A}_0$ from the space of observationally equivalent models in Figure 1. These approaches can be divided into three broad categories: (i) the strong conventional *a priori* identification restrictions on the system matrices, (ii) Bayesian identification using sign priors on the likely effects of certain structural shocks on the endogenous variables, and (iii) statistical identification based on the distributional assumptions about the fundamental innovations. These approaches are listed below in the historical order of appearance in the econometric literature:

- The classical short–run restrictions of Sims (1980), recursive schemes of Christiano, Eichenbaum and Evans (1999) and general non-recursive identification of Sims (1986), Bernanke (1986) and Blanchard and Watson (1986) impose the strong zero *a priori* restrictions on the elements of $\mathbf{A}_0$ or $\mathbf{A}_0^{-1}$, so that the matrix is fully identified in the case of normally distributed structural innovations;

- The long–run restrictions of Blanchard and Quah (1989) are imposed on certain functions of $\mathbf{A}_0$ and the autoregressive matrices $\mathbf{A}_1, \ldots, \mathbf{A}_p$, usually motivated by *ex ante* theoretical considerations, but this identification scheme often needs to be combined with other types of restrictions in order to fully identify the structural model in the normal case;

- The sign restrictions of Faust (1998), Canova and De Nicoló (2002) and Uhlig (2005) use ideas from structural macroeconomic theories on the likely effects of certain structural shocks on the endogenous system variables. This approach is Bayesian in spirit, where the missing part of the model, for which normally distributed structural innovations deliver no identifying information, is recovered by a set of *a priori* assumptions on the impulse response functions; refer to Del Negro and Schorfheide (2011) and Baumeister and Hamilton (2015);

- The new strand of statistical identification literature, exemplified by Lütkepohl and Netšunajev (2014), Herwartz (2016), Kulikov and Netšunajev (2015), Lanne et al. (2017) and the current paper, eliminates all strong *a priori* restrictions on the system matrices $\mathbf{A}_0, \ldots, \mathbf{A}_p$. In this approach the burden of identification is shifted away from the space of system matrices to certain distributional assumptions on the fundamental innovations in the structural model. In the case of independent non–Gaussian innovations, sufficient statistical information is available to identify the structural matrix $\mathbf{A}_0$ without any use of strong priors on its elements.

Although the statistical identification approach to SVAR modelling appears to be eminently attractive because the data is allowed to speak freely on the dynamic effects of fundamental innovations in model (1), even before the economic interpretation (if any) of these shocks becomes clear *a posteriori* (hence the "stargazing" part in the title of the current paper), the technical side of this new approach is not free of certain practical issues. The most fundamental of these issues, which we highlight and discuss in the remaining part of this section, is basically the flip side of the "stargazing" part of the statistical identification methodology: since very little is known *a priori* about the structural matrix $\mathbf{A}_0$ in model (1), one needs to be mindful of the essential multiplicities of the observationally and economically equivalent structural forms corresponding to each particular reduced form model.

The issue of essential multiplicities is not new in the identified VAR literature, but it becomes especially pertinent in the case of statistically identified SVAR models. It has been recently highlighted in this context by Lanne et al. (2010), Kulikov and Netšunajev (2015) and Lanne et al. (2017), where the latter study contains a formal mathematical treatment of the issue. Suppose that we start from a given structural model (1) with a general structural matrix $\mathbf{A}_0$ and pre–multiply it by a given pair of square matrices as follows:

$$\mathbf{PDA}_0\boldsymbol{y}_t = \mathbf{PD}\boldsymbol{k}_0 + \mathbf{PD}\boldsymbol{k}_1 t + \mathbf{PDA}_1\boldsymbol{y}_{t-1} + \ldots + \mathbf{PDA}_p\boldsymbol{y}_{t-p} + \mathbf{PD}\boldsymbol{\epsilon}_t\,, \tag{4}$$

where $\mathbf{P}$ is an $n \times n$ permutation matrix and $\mathbf{D}$ is an $n \times n$ diagonal matrix of plus and minus ones. As one can easily see, this new structural form is observationally equivalent to the original model (1) when transformed into the reduced form (2). Since very little is assumed about the system matrices in statistically identified SVAR models, the two square matrices — $\mathbf{A}_0$ in (1) and $\mathbf{PDA}_0$ in (4) — cannot be *a priori* distinguished.[3] However, the economic interpretation of both models remains the same: their difference is just the order of structural equations (and the individual fundamental innovations in the vector $\boldsymbol{\epsilon}_t$) and the direction of dynamic effects.[4] Importantly, all impulse responses w.r.t. to the fundamental shocks remain intact — the fact that we rely on in Section 3 of the paper for disentangling the order of shocks across the posterior simulations.

Fundamentally, for any given general $n \times n$ structural matrix $\mathbf{A}_0$, there are $n! \cdot 2^n$ essential multiplicities of the structural form (1), which cannot be reduced any further. Since all of these models have the same economic interpretation, their epistemological value from a standpoint of an applied researcher is clearly identical, and the issue of dealing with these multiplicities becomes of a statistical nature.[5] We also emphasise the fact, that the indicated number of

[3] In the language of Lanne et al. (2017), the generalised linear group of non–singular structural matrices, to which $\mathbf{A}_0$ belongs in model (1), is divided into equivalence classes by the actions of $\mathbf{P}$ and $\mathbf{D}$ on this group. Each equivalence class then represents the set of economically similar models. The choice of the best statistical model for a given set of data needs to be confined to one of the equivalence classes, which can be selected arbitrarily.

[4] Since we assume symmetrically distributed shocks in $\boldsymbol{\epsilon}_t$, the direction of dynamic effects in our structural model becomes a matter of normalisation.

[5] Here we rely on the implicit *a priori* assumption that each individual fundamental innovations in the vector $\boldsymbol{\epsilon}_t$ has a distinguishable dynamic effect on the endogenous variables in $\boldsymbol{y}_t$. In order words, each structural innovation in model (1) has a different economic impact and associated interpretation, even if the appropriate theory might not yet be invented! Whether this is so for a given set of observed macroeconomic aggregates, remains to be empirically determined in each application of our methodology.

the essential multiplicities assumes the fixed ordering of observed endogenous variables in $\boldsymbol{y}_t$, otherwise another factor of $n!$ needs to be taken into account.[6] Fortunately, since $\boldsymbol{y}_t$ consists of the observed quantities, the ordering of endogenous variables in the statistically identified SVAR models can be selected arbitrarily and kept fixed in any given empirical application.

Although the issue of essential multiplicities in the statistically identified SVAR models makes no difference from the interpretational perspective in empirical applications, it still needs to be carefully accounted for during the statistical inference stage in our new methodology. We have previously encountered a very similar issue in Kulikov and Netšunajev (2015), where a Bayesian SVAR model with Markov switching volatility in the structural innovations is introduced and applied to a real–world macroeconomic dataset. Section 3 of the current paper builds on these results and gives a full Bayesian statistical inference procedure for the statistically identified SVAR models with independent non–Gaussian innovations, where the essential multiplicities are dealt with after the posterior simulations of the reduced form model parameters are obtained. Our approach allows us to provide a full characterisation of the structural parameter uncertainty, such that the unique fundamental innovations are grouped together according to their dynamic economic impact.

3 Bayesian statistical inference for SVAR models independent non–Gaussian innovations

This section outlines the full Bayesian statistical inference approach to SVAR models with independent non–Gaussian innovations introduced in Section 2. In particular, we describe a fast and efficient clustering–based approach to disentangling the order of structural shocks in different posterior parameter draws.

How does a statistical identification of the mixing matrix $\mathbf{A}_0^{-1}$ in (3) proceeds in a Bayesian SVAR model with independent non–Gaussian structural innovations? It is based on ideas from the independent component analysis (ICA), which is a relatively new method in the technical literature for carrying out "blind" separation of sources from the observed linear mixture. The "blind" part refers to the problem of simultaneous estimation of both the unknown linear mixture matrix $\mathbf{A}_0^{-1}$ and the unknown independent non–Gaussian innovations $\boldsymbol{\epsilon}_t$ from the observed vector of reduced form residuals $\boldsymbol{u}_t$.

The key idea of ICA is that any linear mixture of non–normal structural innovations in (3) is closer to normality by the LLN than any of the individual innovations. Starting form an observed sequence of the reduced form residuals $\{\boldsymbol{u}_t : 1 \leq t \leq T\}$ with (almost) normal multivariate densities, one obtains an $n \times n$ de–mixing matrix $\mathbf{W}$ s.t. the sequence of $\{\tilde{\boldsymbol{\epsilon}}_t : 1 \leq$

[6] Indeed, the order of the elements in $\boldsymbol{y}_t$ across the models (1) and (4) remains unchanged. Notice the difference with the classical triangular identification of the structural matrix $\mathbf{A}_0$: it relies on *a priori* linkage between zeros in $\mathbf{A}_0$ and the order of observables in $\boldsymbol{y}_t$, whereby any re–ordering of the structural equations and shocks in the vector $\boldsymbol{\epsilon}_t$ breaks the triangular form of the structural matrix.

$t \leq T\}$, where:

$$\tilde{\boldsymbol{\epsilon}}_t = \mathbf{W}\boldsymbol{u}_t \quad \text{for all } 1 \leq t \leq T\,,$$

is mutually independent and non–normal. The crucial assumption of the ICA is mutual independence and non–normality of the elements of $\tilde{\boldsymbol{\epsilon}}_t$ (only one of them can optionally be normal), and the search for $\mathbf{W}$ proceeds by testing non–normality of the pre–whitened sequence of the reduced form residuals using higher–order cumulant–based contrast functions; see Hyvärinen and Oja (2000) and Hyvärinen (2013).

How close to each other are the structural matrix $\mathbf{A}_0$ on one hand and the estimated de–mixing $\mathbf{W}$ on the other? As emphasised by Hyvärinen and Oja (2000) in the technical literature and by Lanne et al. (2010) in the context of statistical identification of SVAR models, the matrix $\mathbf{A}_0$ can only be identified up to the permutation order and size and sign normalisation of its rows. As explained in Section 2 of the paper, although the space of observationally and economically equivalent models in the diagrammatic language of Figure 1 is reduced to a countable subset of $n! \cdot 2^n$ models, one still needs to make the final choice by selecting one of them. Fortunately, this choice involves only the order of structural equations (and the fundamental shocks in $\boldsymbol{\epsilon}_t$) in (1) for a given ordering of the observables in $\boldsymbol{y}_t$, together with the sign and scale normalisation of the innovations in $\boldsymbol{\epsilon}_t$, and therefore does not substantively alter or affect in any other way the dynamic impact of each individual structural innovation on the set of endogenous variables in the model. This choice is therefore innocuous from the applied macroeconomic perspective.

When a single point estimate of $\mathbf{A}_0$ is sought in the maximum–likelihood based inferential context, this problem can be dealt with by carefully selecting the optimisation space such that the order, size and signs of the row vectors in the structural matrix $\mathbf{A}_0$ remain the same throughout the likelihood function maximisation procedure. This is equivalent to an *a priori* choice of one of the observationally equivalent models, but since only a single point parameter estimate is needed, no other statistical issues arise in this context.

However, in the Bayesian statistical context, a whole posterior distribution of the structural matrix $\mathbf{A}_0$ has to be evaluated by one of the numerical posterior simulation procedures, and there is no guarantee that each posterior draw of model parameters would result in exactly the same structural innovations order as all the other posterior draws. This merely reflects the fact that, in contrast to a single point estimation in the maximum–likelihood context, where a particular choice of the structural matrix has to be done only once, the identification problem is compounded in the Bayesian context by the need to evaluate an entire distribution of $\mathbf{A}_0$ in order to be able to fully characterise the posterior uncertainty of this model parameter. Notice that this is not a conceptual identification issue, merely a practical difficulty impairing the full statistical identification of model (1) in the Bayesian statistical context.

Therefore, a big practical difficulty for the full Bayesian inference on the structural form parameters of model (1) is that for every posterior draw of the reduced form residuals from the VAR model (2) we are facing a potentially different order of the rows in the structural system (the size of structural innovations is usually normalised to unity by assuming that $\mathbb{E}\,\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t' = \mathbf{I}$). Hence, all posterior impulse response functions and other posterior statistics are going to be mixed up across different draws from the posterior distribution.

Fortunately, as in Kulikov and Netšunajev (2015), we already have a solution — a clustering–based re–ordering of the posterior $\mathbf{W}$ matrix draws. It is based on a simple and powerful idea: although the order of rows in the system matrix $\mathbf{A}_0$ across differen posterior draws is arbitrary (there are $n!$ different combinations), all structural innovations in the original model (1) and any of its row–permuted versions (4) retain their unique economic interpretations reflected in their specific dynamic impact on the endogenous system variables.

In Kulikov and Netšunajev (2015), the unique structural shocks were disentangled from each other by carrying out the clustering procedure on the posterior impulse response function draws. In this paper, a similar clustering–based approach is used directly on the posterior draws from the $\{\tilde{\boldsymbol{\epsilon}}_t : 1 \leq t \leq T\}$ sequence: each of these draws is an estimate the original sequence of non–normal independent structural innovations $\{\boldsymbol{\epsilon}_t : 1 \leq t \leq T\}$, but potentially in a different order and with a different sign.

By using a cross–correlations based procedure on the entire set of posterior draws of $\{\tilde{\boldsymbol{\epsilon}}_t : 1 \leq t \leq T\}$, a very fast column and sign re–ordering of the posterior $\mathbf{W}$ matrix draws becomes feasible and the final system identification of our SVAR model possible.

To sum up, the new ICA–based statistical identification of Bayesian SVAR models proceeds in the following steps:

1. Any Bayesian statistical inference starts with priors: here the priors are imposed on the reduced form VAR parameters, but they do have an indirect impact on the structural matrices as well;

2. Using the priors and the data, generate a sufficient number of posterior draws of the reduced form VAR parameters, including the sequence of reduced form residuals $\{\boldsymbol{u}_t : 1 \leq t \leq T\}$;

3. Run your favourite ICA algorithm (we use FastICA by Hyvärinen and Oja (2000)[7]) on each posterior draw of $\{\boldsymbol{u}_t : 1 \leq t \leq T\}$, getting a sequence of posterior draws of de–mixing matrices $\mathbf{W}$ and innovations $\{\tilde{\boldsymbol{\epsilon}}_t : 1 \leq t \leq T\}$;

4. Run our clustering–based procedure on the posterior draws of $\{\tilde{\boldsymbol{\epsilon}}_t : 1 \leq t \leq T\}$ and re–order the rows and normalise the signs of $\mathbf{W}$ accordingly, finally getting the posterior draws of the system matrix $\mathbf{A}_0$;

5. Compute any other desired posterior statistics.

[7] A fast and efficient MATLAB® implementation of this numerical algorithm is made available at `research.ics.aalto.fi/ica/fastica` by its original authors.

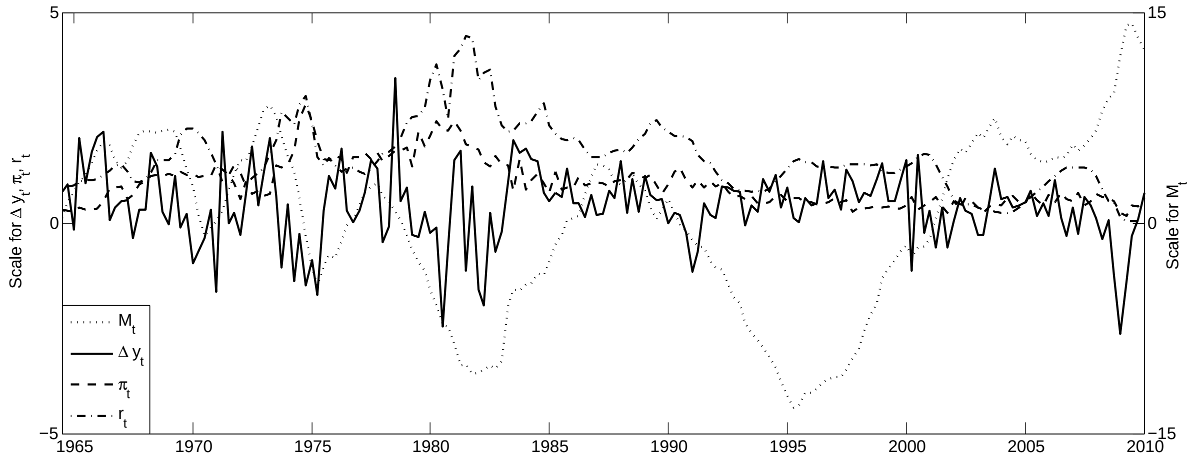

Figure 2: The US macroeconomic data series from 1964Q2 to 2009Q4

4 Empirical application to the US macroeconomic dynamics

To illustrate the new ICA–based statistical identification of SVARs, we utilise a quarterly US macroeconomic dataset originally used for the illustration of the heteroscedasticity–based SVAR identification in Kulikov and Netšunajev (2015). This dataset provides a good benchmarking target for the new ICA–based approach, since in the previous study we have successfully identified the effects of four structural innovations in this data and have found their plausible macroeconomic sources. In this section we are aiming to carry out a similar statistical identification exercise using the new ICA–based identification technique outlined in Sections 2 and 3 of the paper.

The US macroeconomic data is quarterly, seasonally adjusted, covering time period from 1964Q2 to 2009Q4, supplied by the Federal Reserve Bank of St. Louis FRED database.[8] Per capita aggregates are computed using the US civilian non-institutional population aged from 16 years up. Figure 2 displays the data:

- Output growth rate (Δy_t) is a scaled quarter-on-quarter log real GDP per capita change;
- Inflation rate (π_t) is a scaled quarter-on-quarter log change of personal consumption expenditures core price index;
- Real money balance (M_t) is a sum of de-trended log inverse money velocity, calculated as a ratio of the quarterly sweep-adjusted M2 money stock and quarterly nominal GDP, and log real output per capita;
- Monetary policy interest rate (r_t) is a quarterly average of the federal funds rate.

It is the same dataset that was used previously by Kulikov and Netšunajev (2015) in the Markov switching Bayesian SVAR context. We more or less know what to expect from these data — in this paper we only aim to benchmark our new ICA–based statistical identification procedure for SVAR models against the previously documented results.

[8]All data series are downloaded from `research.stlouisfed.org/fred2`

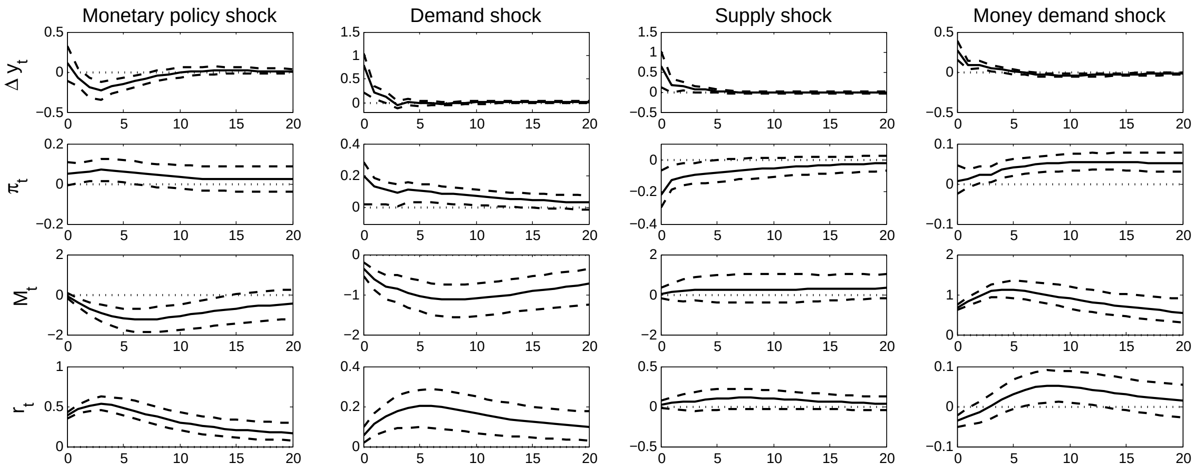

Figure 3: Posterior impulse response functions and 68% credible sets for SVAR model identified using estimated Markov switching volatility regimes

The previous results are summarised in Figure 3, where the posterior impulse response functions with 68% credible sets are plotted for a set of four structural shocks, which are statistically identified using the estimated volatility regime shifts in the US macroeconomic dynamics over the sample period. The most interesting finding documented in Kulikov and Netšunajev (2015) pertains to relatively clear–cut economic interpretation of these four shocks, that are found in the data without the help of strong conventional *a priori* identification restrictions on the system matrices. In particular, we were able to identify the short–run impacts and impulse responses of four structural shocks, which we labeled a “monetary policy shock”, a “money demand shock”, an “aggregate demand shock” and an “aggregate supply shock”. We also found a robust and well–pronounced “price puzzle” in response to a “monetary policy shock”, while the “money multiplier shock” induces a strong protracted response of the US real output and prices in the sample data.

In order to fully disentangle these shocks during the posterior simulation procedure, we also used a clustering procedure, similar to that introduced in Section 3, but applied instead to the set of estimated posterior impulse responses from the Gibbs sampler. This clustering method delivered enough statistical evidence to confidently discern the two middle shocks in Figure 3, for which the original Gibbs sampler delivered the intermixed order across different posterior draws.

In this paper, we use the same US macroeconomic dataset to apply and test the effectiveness of our full Bayesian statistical inference procedure for the SVAR models with independent non–Gaussian innovations outlined in Section 3. We start with a reduced form VAR model with independent normal–Wishart priors on the reduced form parameters $\mathbf{\Phi}_1, \ldots, \mathbf{\Phi}_p$ and the variance–covariance matrix of the reduced form residuals $\boldsymbol{u}_t$. In this case the structural model remains essentially unidentified, since we do not assume any of the strong conventional *a priori* restrictions on the structural form parameters. Using the four steps of our full Bayesian

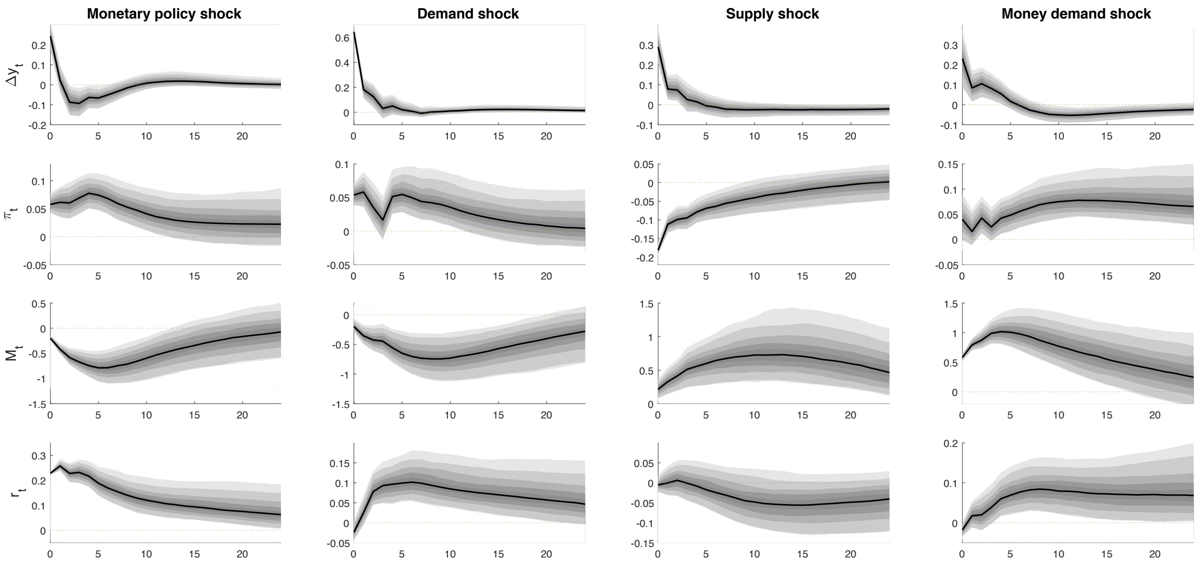

Figure 4: Posterior impulse response functions and 96% credible sets for SVAR model with independent non–Gaussian innovations identified using the ICA–based approach

statistical inference procedure from Section 3, which includes the use of independent component analysis on each posterior draw of the reduced form residuals followed by our clustering–based procedure for disentangling the order of the structural innovations, we are able to generate a sufficient number of posterior draws of $\mathbf{\Phi}_1, \ldots, \mathbf{\Phi}_p$ and the associated system matrix $\mathbf{A}_0$ to compute the posterior impulse response functions in Figure 4.

Notice that in contrast to Kulikov and Netšunajev (2015), where only the two middle shocks in Figure 3 had to be re–ordered from the Gibbs sampler posterior output, in this paper the order of all four shocks remained intermixed in the set of original posterior draws. The computational burden of the clustering–based procedure is therefore proportional to the factorial of the size of the estimated VAR model. However, in terms of the total computation time for the full posterior simulation algorithm in Section 3, the clustering–based shock re–ordering procedure takes up only a small share. The biggest computation burden is imposed by the FastICA algorithm during the estimation of the de–mixing matrix $\mathbf{W}$, prior to its re–ordering to the final posterior draw of the structural matrix $\mathbf{A}_0$.

In terms of the posterior impulse response functions in Figure 4, we observe much sharper statistical results compared to the Markov switching volatility approach in Figure 3. Notices that we display 96% credible sets around the median impulse responses in Figure 4, compared to the 68% credible sets in Figure 3. Nevertheless, the posterior uncertainty for our empirical SVAR model with independent non–Gaussian innovations appears to be considerably narrower than in the Markov switching volatility–identified SVAR case in Kulikov and Netšunajev (2015). We interpret this finding as a clear indication that the non–Gaussian innovations in our statistical model deliver sufficient information for the sharp identification of the structural matrix $\mathbf{A}_0$ and other model parameters.

This can also be seen by the plot of prior and posterior distributions of the elements of

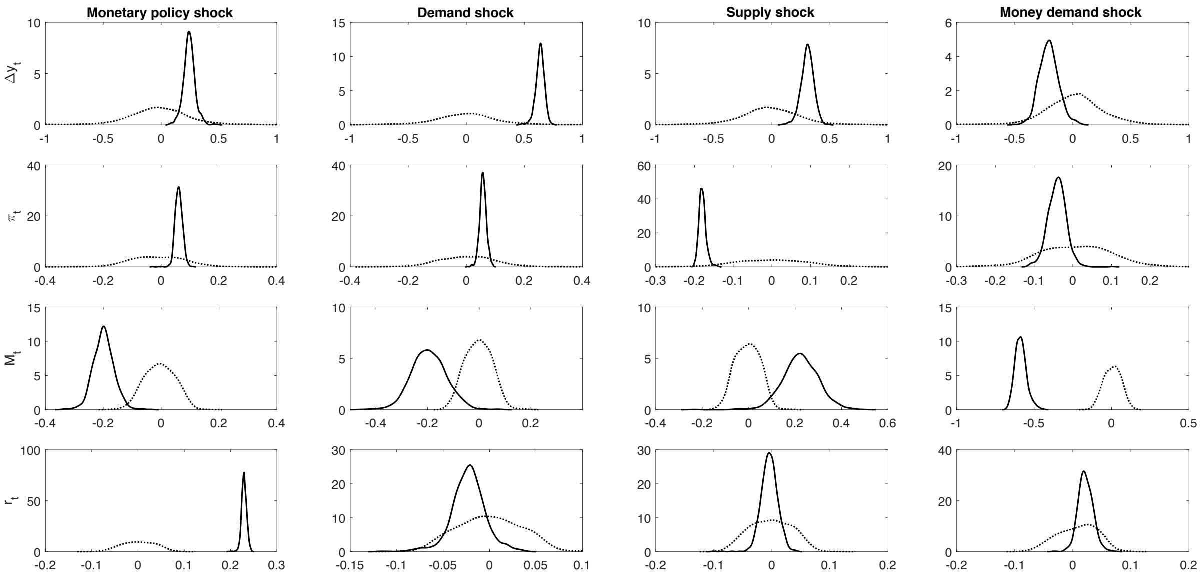

Figure 5: Prior and posterior distributions of $\mathbf{A}_0^{-1}$ for SVAR model with independent non–Gaussian innovations identified using the ICA–based approach

the inverse structural matrix, $\mathbf{A}_0^{-1}$, in our statistical model, displayed in Figure 5. All contemporaneous effects of four structural shocks appear to be well–pronounced, with only a few of them having a non–negligible posterior probability of a zero on–impact effect on the endogenous model variables. The data also appears to be fairy informative about the structural matrix in our statistical model, which we judge here by the relatively large changes between the non–informative priors and the estimated posterior distributions of all elements of the $\mathbf{A}_0^{-1}$ matrix in Figure 5.

The economic interpretation of our empirical findings in Figures 4 and 5 appears somewhat different from the Markov switching volatility–identified SVAR model in Kulikov and Netšunajev (2015). The most striking difference concerns the "monetary policy shock" in the first column of Figures 3 and 4. While the Markov switching volatility–based statistical identification delivers a comfortably interpretable results of a policy rate change on the real and monetary sides of the US economy, with a slightly upsetting finding of a pronounced "price puzzle" in response to the monetary policy intervention[9], the nature of the endogenous model responses to the "monetary policy shock" in the first column of Figure 4 appears to be even less conventional. In particular, we find a high posterior probability and pronounced real output growth response to our statistically identified "monetary policy shock", which at the very short run horizon is positive and runs opposite to the conventional intuition about the monetary policy. We provisionally dub this effect a "real output puzzle", in the same vein as the familiar "price puzzle", which is also present in our statistical model in Figures 4 and 5. At the juncture, there is no good explanation of the newly identified "real output puzzle", but we conjecture that an economic intuition behind it might be closely linked to that of the "price

[9]We are aware of the need to include an extra information variable into the system, such as a commodity price index.

puzzle" — due to an insufficient conditioning set in our relatively small–scale SVAR model, the first shock might in fact be a combination of several conceptually different innovations that all happen to strongly influence the monetary policy rate on impact in our statistical model.

Another pretty large difference between the two sets of results is related to the pronounced movement of the real money aggregate in response to the statistically identified "supply shock" in our empirical model. However, it appears to be in line with the economic intuition, as we observe a strong negative effect on the prices in response to the "supply shock", while the monetary policy appears to be the passive mode. In this case, we expect the real money balances to increase. This is an example, where our previous Markov switching volatility–identified SVAR model did not succeed to deliver a strong identification of this effect, while the new approach, based on the statistical information contained in independent non–Gaussian innovations, is clearly superior in this respect.

To sum up, we have compared the performance of our new ICA–based statistical identification technique for SVAR models with the previous results in Kulikov and Netšunajev (2015), which relied on the volatility shifts in the conditionally normally distributed innovations to identify and estimate the effects of four structural shocks in the US macroeconomic data series over the period from 1964Q2 to 2009Q4. We find that the new model identification approach delivers sharper statistical results and narrower posterior confidence sets around the estimated impulse responses. While all four previously identified structural shocks are also picked up by the new model, we find that some previously wide uncertainty bands have narrowed down to the point, where we can see some new statistically non–negligible effects with high posterior probability. In particular, we document a "real output puzzle" in response to the statistically identified "monetary policy shock", and we clearly see the previously uncertain response of the real money aggregate in the "supply shock" scenario.

5 Conclusion

In this paper, we propose a new ICA–based statistical identification procedure for SVAR models with independent non–Gaussian structural innovations, along the lines of Herwartz (2016) and Lanne et al. (2017). We also work out details of the fast and efficient Bayesian posterior sampling algorithm that fully identifies and disentangles the intermixed structural shocks in different posterior parameter draws. The most attractive feature of the new statistical identification method is that it allows us to use additional statistical information available in non–normal shocks to identify all elements of the structural matrix $\mathbf{A}_0$ in the model without any use of the strong conventional *a priori* identifying restrictions.

The new procedure is validated using the US macroeconomic data series, where the nature of different shocks is empirically examined. For this validation exercise, we used the same dataset and previously documented results in Kulikov and Netšunajev (2015), which relied on the volatility shifts in the conditionally normally distributed innovations to identify and estimate the effects of four structural shocks in the US macroeconomic data series over the period from 1964Q2 to 2009Q4. As in the previous study, we identify the short–run impacts and impulse responses of four structural shocks, which we label a "monetary policy shock", a "money demand shock", an "aggregate demand shock" and an "aggregate supply shock". We also find that the new model identification approach delivers sharper statistical results and narrower posterior confidence sets around the estimated impulse responses. While all four previously identified structural shocks are also picked up by the new model, we find that some previously wide uncertainty bands have narrowed down to the point, where we can see some new statistically non–negligible effects with high posterior probability. In particular, we document a "real output puzzle" in response to the statistically identified "monetary policy shock", and we clearly see the previously uncertain response of the real money aggregate in the "supply shock" scenario.

What's next? How about identifying more structural innovations from the data than there are dimensions available in VAR model?

References

Amisano, G. and Giannini, C. (1997). *Topics in Structural VAR Econometrics*, 2nd edn, Springer, Berlin.

Baumeister, C. and Hamilton, J. D. (2015). Sign restrictions, structural vector autoregressions, and useful prior information, *Econometrica* **83**: 1963 – 1999.

Bernanke, B. S. (1986). Alternative explanations of the money-income correlation, *Carnegie-rochester conference series on public policy*, Vol. 25, Elsevier, pp. 49–99.

Blanchard, O. J. and Watson, M. W. (1986). Are business cycles all alike?, *The American business cycle: Continuity and change*, University of Chicago Press, pp. 123–180.

Blanchard, O. and Quah, D. (1989). The dynamic effects of aggregate demand and supply disturbances, *The American Economic Review* **79**(4): 655–673.

Canova, F. and De Nicoló, G. (2002). Monetary disturbances matter for business fluctuations in the G-7, *Journal of Monetary Economics* **49**: 1131–1159.

Christiano, L. J., Eichenbaum, M. and Evans, C. L. (1999). Monetary policy shocks: What have we learned and to what end?, Vol. 1, Part A of *Handbook of Macroeconomics*, Elsevier, pp. 65 – 148.

Comon, P. and Jutten, C. (2010). *Handbook of blind source separation, independent component analysis and applications*, Academic Press.

Darmois, G. (1953). Analyse générale des liaisons stochastiques: etude particulière de l'analyse factorielle linéaire, *Review of the International Statistical Institute* **21**: 2 – 8.

Del Negro, M. and Schorfheide, F. (2011). Bayesian macroeconometrics, *in* J. Geweke, G. Koop and H. van Dijk (eds), *The Oxford Handbook of Bayesian Econometrics*, Oxford University Press, pp. 293 – 389.

Faust, J. (1998). The robustness of identified VAR conclusions about money, *Carnegie Rochester Conference Series on Public Policy 49*.

Hamilton, J. D. (1994). *Time Series Analysis*, Princeton University Press, USA.

Herwartz, H. (2016). Structural VAR modelling with independent shocks, *Preprint*.

Herwartz, H. and Lütkepohl, H. (2014). Structural vector autoregressions with markov switching: Combining conventional with statistical identification of shocks, *Journal of Econometrics* **183**(1): 104 – 116.

Hyvärinen, A. (2013). Independent component analysis: Recent advances, *Philosophical Transactions of the Royal Society A* **371**.

Hyvärinen, A. and Oja, E. (2000). Independent component analysis: Algorithms and applications, *Neural Networks* **13**: 411 – 430.

Hyvärinen, A., Zhang, K., Shimizu, S. and Hoyer, P. O. (2010). Estimation of a structural vector autoregression model using non–Gaussianity, *Journal of Machine Learning Research* **11**: 1709 – 1731.

Kulikov, D. and Netšunajev, A. (2015). Identifying shocks in structural VAR models via heteroscedasticity: a Bayesian approach, *Working Paper Series 8*, Eesti Pank.

Lanne, M. and Lütkepohl, H. (2008). Identifying monetary policy shocks via changes in volatility, *Journal of Money, Credit and Banking* **40**: 1131–1149.

Lanne, M., Lütkepohl, H. and Maciejowska, K. (2010). Structural vector autoregressions with Markov switching, *Journal of Economic Dynamics and Control* **34**: 121–131.

Lanne, M., Meitz, M. and Saikkonen, P. (2017). Identification and estimation of non–Gaussian structural vector autoregressions, *Journal of Econometrics* **196**(1): 288 – 304.

Leeper, E. M., Sims, C. A. and Zha, T. (1996). What does monetary policy do?, *in* R. E. Hall and B. S. Bernanke (eds), *Brooking Papers on Economic Activity*, Brooking Institution Press, pp. 1 – 78.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*, Springer-Verlag, Berlin.

Lütkepohl, H. and Netšunajev, A. (2014). Disentangling demand and supply shocks in the crude oil market: How to check sign restrictions in structural VARs, *Journal of Applied Econometrics* **29**(3): 479–496.

Moneta, A., Entner, D., Hoyer, P. O. and Coad, A. (2013). Causal inference by independent component analysis: Theory and applications, *Oxford Bulletin of Economics and Statistics* **75**: 705 – 730.

Rigobon, R. (2003). Identification through heteroskedasticity, *Review of Economics and Statistics* **85**: 777–792.

Rubio-Ramírez, J. F., Waggoner, D. F. and Zha, T. (2010). Structural vector autoregressions: theory of identification and algorithms for inference, *Review of Economic Studies* **77**: 665 – 696.

Sims, C. (1980). Macroeconomics and reality, *Econometrica* **48**(1): 1–48.

Sims, C. A. (1986). Are forecasting models usable for policy analysis?, *Federal Reserve Bank of Minneapolis Quarterly Review* **10**(1): 2–16.

Uhlig, H. (2005). What are the effects of monetary policy on output? Results from an agnostic identification procedure, *Journal of Monetary Economics* **52**(2): 381–419.