Alma Mater Studiorum - Università di Bologna

DOTTORATO DI RICERCA IN
Scienze Statistiche
XXIX Ciclo

Settore Concorsuale di afferenza: 13/A5

Settore Scientifico disciplinare: SECS-P/05

Identification and estimation of Structural VAR models with mixed frequency data: a moment-based approach

Presentata da
Silvia Emili

Coordinatore Dottorato
Prof.ssa Alessandra Luati

Relatore
Prof. Luca Fanelli

Esame finale anno 2017

Contents

List of Figures

List of Tables

Summary

The frequency at which institutional and government economic agencies take decisions is not necessarily the one at which data are published. The mismatch between the "*time*" process which characterizes agent's decisions and data availability is very common in economics and finance. Even if a non negligible part of economic information is available at low frequencies, making decisions at frequency higher than data availability is an usual approach.
A classical example is represented by the analysis of the co-movement between output growth, interest rate and inflation. Usually, the variable chosen as proxy of output growth is the Gross Domestic Product (henceforth, GDP), which is officially published at quarterly frequency. On the other hand, it is usual to proxy the interest rate and inflation with Federal Fund Rate (FFR) and the Consumer Price Index (CPI), respectively. In this situation, the classical approach (henceforth, *naive approach*) corresponds to aggregate the (high frequency) monthly variables until all the data can be treated as quarterly variables, i.e. the data are reported at the same low frequency. This procedure embodies some critical consequences in term of estimation bias and interpretation of results. By this way, the econometric literature analysed these consequences, both in the univariate and multivariate framework. Among the others, Christiano and Eichenbaum (1987) and Marcellino (1999) furnish a wide perspective of estimation bias due to temporal aggregation. In particular, Christiano and Eichenbaum (1987) quantify the error due to mispecification of the temporal interval at which economic agents make decisions. Marcellino (1999) identifies and summarizes which properties of the time series processes remain invariant and which vary after temporal aggregation. The most intuitive example is represented by Impulse Response functions (IRFs). The structural analysis, and in particular IRFs, belong to the first category: the aggregation of impulse responses cause naturally a bias in the aggregated results, and hence a wrong interpretation of a variable's reaction to a shock in another variable.
In literature, an interesting experiment about temporal aggregation bias in structural analysis, is proposed by Cunningham and Hardouvelis (1982). The authors investigate how temporal aggregation affects the role of liquidity effect on the mechanisms of transmission of monetary policy effects. Cunningham and Hardouvelis

(1982) find that the estimated responses substantially change in their magnitude when they consider to work with variables aggregated at the same low frequency. The loss of information and the wrong interpretation of results obtained with the *naive approach*, can be seen as the consequence of a mispecification of the co-movements between the variables in the dataset. In last years the literature pays particular attention to develop econometric models able to investigate all the information in the mixed frequency dataset. In this direction, incisive approaches have been proposed by Mariano and Murasawa (2003, 2010), Clements and Galvão (2008) and Ghysels (2012). These methodologies are characterized by the idea of exploit the high frequency information in order to reduce the mispecification of the high frequency (real) joint process.
In Chapter 1 we consider one of the most attractive tool provided by this field of research: Mixed Frequency Vector Autoregressive processes (MF-VARs). A mixed frequency VAR can be defined as a VAR model that, in a mixed frequency dataset, investigates the dynamics between (i) the highest frequency (observed) variables and (ii) the high frequency latent processes underlying the observed low frequency series. Nowadays, the MF-VARs literature appears as a growing econometric framework. As shown below in Chapter 1, numerous are the contributions provided by the researchers. Among the others, Camacho (2013) and Foroni, Guerin and Marcellino (2014) specify a Markov-switching dynamics in the MF-VAR, Marcellino, Porqueddu and Venditti (2015) introduce MF-VAR with stochastic volatility, Marcellino and Sivec (2015) pay attention to MF-FAVAR models.
After the introduction of MF-VAR models, some works have been provided in order to quantify analytically and/or computationally the loss of information generated by

- the use of mixed frequency data instead of working with high frequency variables,
- the use of low frequencies opposite to mixed frequency variables.

Koelbl, Braumann, Felsenstein and Deistler (2016) propose to measure this loss of information resorting either to the asymptotic variance of the estimates, or to the one-step ahead prediction error variances. The results are obtained comparing the solutions obtained with standard estimation methods for high frequency and low frequency processes, and the results obtained with the Extended Yule-Walker (XYW) equations of Chen and Zadrozny (1998) for the mixed frequency case. Differently, Foroni and Marcellino (2014) refer to MF-VAR models with state space representation (see Mariano and Murasawa (2003, 2010)). They assume that the actual data generating process (henceforth, DGP) is at (high) monthly frequency, but the econometrician can only observe the aggregated quarterly realizations for

some variables. Foroni and Marcellino evaluate the differences among the results obtained with the aggregated (quarterly) solution and the mixed frequency case. Specifically, the authors consider a simplified specification of the new Keynesian dynamic stochastic general equilibrium (DSGE) model analysed by Clarida, Galì and Gertler (2000), and a second version obtained inserting in the dynamic Investment and Saving (DIS) equation one lag of order one of the output growth, *à la* Fuhrer and Rudebusch (2004). In both the examples, Foroni and Marcellino find a mismatch and a loss of information between the naive approach and the aggregated solution, once assumed that the real DGP is at high frequency. The authors highlights also a loss of identification in the second DSGE model due to high non-linearity provided by temporal aggregation.
Foroni and Marcellino demonstrate that, considering a suitable estimation procedure, the analysis of mixed frequency data alleviate the classical bias due to temporal aggregation. The estimation procedure proposed by Mariano and Murasawa (2003, 2010) and chosen by Foroni and Marcellino (2014), is nowadays one of the most implemented approach in dealing with mixed frequency data. This approach consists in referring to the state space representation of the MF-VAR.
In this thesis, in Chapter 1 we provide a general treatment of the MF-VAR literature. We identify two classes of empirical and economical motivations behind the research activity about mixed frequency data: temporal aggregation/disaggregation and nowcasting. The attention of the researchers on mixed frequency data exponentially enhanced with the phenomenon of nowcasting. The *classical* definition of this forecasting issue is provided by Banbura, Giannone, Modugno and Reichlin (2013): "[...]*Now-casting is defined as the prediction of the present, the very near future and the very recent past. The term is a contraction for now and forecasting and has been used for a long-time in meteorology and recently also in economics*".
In order to model different economic phenomena, in the MF-VAR literature, many applications and developments of MF-VAR models have been provided in the last decade. After a briefly review of the developments provided in the literature, we pay particular attention to the Structural analysis of MF-VARs.
In Chapter 2, we provide a novel approach for the estimation of MS-SVARs. This technique appears particular useful in those situations in which the researcher is interested in the analysis of the high frequency process, but the high information in the data is limited or null. The proposed estimation approach investigate the mapping between the high frequency process and the aggregated low frequency counterpart. In the discussion we consider different aggregation schemes and we generalize the estimation approach to different identification structural schemes. We compare the results obtained with the novel approach to the state space procedure with different Monte Carlo experiments. Moreover, we propose a generalization to higher order VAR to the estimation of the IRFs of the high frequency

process. The procedure involves the Impulse Response Function Matching Estimator proposed by Rotemberg and Woodford (1997). Following Guerron-Quintana, Inoue and Kilian (2017), we adapt this techniques to our objective, and we verify the reliability of the results with a Monte Carlo simulation. The Empirical illustration is implemented referring to this last proposal.

Chapter 1

Mixed frequency VARs: a new generation of econometric models.

1.1 Introduction

Working with economic time series sampled at different frequency, is one of the most common situation in empirical analyses. The technologies developments and a growing interest in the effects of the co-movements between financial markets, national economies, and the decision processes of economic agents, provide a big quantity of information available for empirical analyses. As widely depicted in literature, the classical approach (henceforth, *naive* approach) of aggregating the high frequency variables in a mixed frequency dataset, until all the data present the same low frequency, can lead to different identification, estimation and interpretation problems. Forty years of literature about temporal aggregation have shown the limits and the shortfalls of this approach[1]. In particular, the researchers emphasized the unreliability of the results and their interpretation, as a consequence of the misspecification of the co-movements between variables sampled at different frequency, and then aggregated at the same low frequency.
To mitigate the phenomenon of aggregation bias, in the following years, the literature gradually focuses on the specification of econometric models able to consider all the available information in the data. In the latest '80 and '90, Zadrozny (1988) and Chen and Zadrozny (1998) investigate continuous-time autoregressive models with mixed frequency variables. They consider the problem of dealing with mixed frequency data in which the low frequency variables are generated by an underlying latent high frequency process[2]. However, the attention of the re-

[1] Cunningham and Hardouvelis (1982), Christiano and Eichenbaum (1987), Marcellino (1999), among others

[2] Chen and Zadrozny (1998) consider a classical representation of VAR processes with missing data and propose an Extended Yule-Walker method for the estimation of the VAR parameters.

searchers on mixed frequency data exponentially enhanced only fifteen years later (approximately), with the concept of nowcasting. Nowadays, this phenomenon is introduced in the literature as an attractive, meaningful and powerful framework. The *classical* definition of this forecasting issue is provided by Banbura, Giannone, Modugno and Reichlin (2013): "[...]*Now-casting is defined as the prediction of the present, the very near future and the very recent past. The term is a contraction for now and forecasting and has been used for a long-time in meteorology and recently also in economics*". In particular, nowcasting is usually presented as a fast, ever-expanding phenomenon due to the attention of national and international central banks: from Kuzin, Marcellino and Schumacher (2013), "[...] *decision-makers in policy institutions typically face such data irregularities in their everyday business of assessing the current state of the economy*".
Tracking the swings of the economy furnishes relevant information to government and institutional agencies, allowing those agents to reactive political and economical reactions. The last global crisis is representative of this current need for more timely detection of certain situations of real risk for the economic equilibrium.
Within the set of econometric instruments provided by the researchers in the nowcasting framework, Banbura, Giannone, Modugno and Reichlin (2013) distinguish between: (i) models with temporal aggregation (for a survey, see Marcellino (1999)); (ii) joint models in state space representation (factor models and mixed frequency VARs) as in Mariano and Murasawa (2003, 2010), Giannone, Reichlin and Small (2008), Schorfheide and Song (2015); (iii) updating frameworks (see for example Banbura and Modugno (2014) and Banbura and Rünstler (2011)); (iv) partial models (bridge and MIDAS-type equations) as in Baffigi, Golinelli and Parigi (2004), Ghysels, Sinko and Valkanov (2006) and Clements and Galvão (2008, 2009). In the following years, these strands of the literature are not infrequently encountered and merged, furnishing a wide range of approaches to nowcasting exercises.
Generally, within the classes identified above, the choice on the specification of an econometric model for mixed frequency data, moves between (1) the specification of a joint model for the variable of interest and for the predictors (full system approaches, i.e. MF-VARs)[3], or (2) a model uniquely specified to model the low frequency variable of interest as dependent on mixed frequency variables (partial models, i.e. MIDAS equations)[4].

The result is an optimal three-step (linear) instrumental variables procedure.

[3]Next to the strand of the literature of MF-VAR, another important direction is represented by the general idea of summarizing a large dataset of mixed frequency variables, in a small number of factors or coincident indicators, which are obtained for a unique frequency (for all, see Stock and Watson (2002)).

[4]Kuzin, Marcellino and Schumacher (2011) provide a theoretical and empirical comparison, based also on the nowcasting performances, between MIDAS equations and MF-VARs. The

In the class of full system methods, the most incisive approach is presented by Mariano and Murasawa (2003, 2010). In their seminal papers, the authors are driven by the idea of alleviating two shortcomings in the definition of coincident indexes of the business cycle: the loss of information due to the non-inclusion of quarterly variables, and the shortfall of an economic interpretation. Mariano and Murasawa consider the state space representation with the aim of treating low frequency data (quarterly) as related to a latent process, representative of the high (monthly) levels[5].
If with a MF-VARs the researcher models the latent high frequency process that define the co-movements between the variables of the analyses, Ghysels (2016) introduces a new perspective of the problem, more in line with the MIDAS literature. The author defines a stacked MF-VAR in which the referring time interval of the econometric model is the lowest frequency in the data. The general idea consists of stack the high and the low observations referring to the same low frequency period. By this way, the author investigates different scenarios for the formulation of the stacked endogenous vector, the structural analysis of the new specification, and highlights the major differences between the mixed frequency VAR with a state space representation and his proposal.
In the next sections we consider both the specifications and we analyse the main aspects of each approach. In particular, in Section 1.2, we describe the methods used in the literature to specify and estimate a VAR for mixed frequency data. In Section 1.4, we consider the literature of MF-VARs, related empirical applications and extensions, paying particular attention, in Section 1.5, to mixed frequency Structural Vector Autoregressive processes. In Section 1.6 we report our considerations and conclusions.

1.2 MF-VAR models

A mixed frequency vector autoregressive model (hereafter, MF-VAR), is a vector autoregressive process specified to deal with data sampled at different frequencies. Let $y_{1,t}, t = 1, \ldots, T$, be the N_1-variate processes of quarterly (low frequency) variables, observed only in the third month of the reference quarter. By the same way, let $y_{2,t}, t = 1, \ldots, T$, be the N_2-variate monthly (high frequency) process sampled $m = 3$ times during each reference quarter. Consider the vector $y_t = (y'_{1,t}, y'_{2,t})'$, of dimension $N_1 + N_2 = N$. The MF-VAR investigates the dynamic interaction

authors find that the two approaches could be seen as complementary tools, instead of substitutes. Aastveit, Gerdrup, Jore and Thorsrud (2014) present a combination approach of 224 models obtained from bridge equations, factor models and mixed-frequency VARs.

[5]Similar approaches have been provided by Schorfheide and Song (2015) and Giannone, Reichlin and Small (2008).

between the highest frequency (observed) variables $y_{2,t}$ and the high frequency latent processes $y^*_{1,t}$ underlying the observed low frequency series $y_{1,t}$.
The most implemented MF-VAR specification is provided by Mariano and Murasawa (2003, 2010). It requires the state space representation of the model as a tool for the joint investigation of (i) the latent processes $y^*_{1,t}$ underlying the observed low frequency series $y_{1,t}$ and (ii) the monthly Data Generating Process (hereafter, DGP) $y^*_t = (y^{*'}_{1,t}, y'_{2,t})'$ of $y_t = (y'_{1,t}, y'_{2,t})'$. The MF-VAR(p) is defined as

$$y^*_t = \Phi_1 y^*_{t-1} + \Phi_2 y^*_{t-2} + \cdots + \Phi_p y^*_{t-p} + \varepsilon_t, \quad \varepsilon_t \sim iid(0, \Sigma), \quad t = 1, \ldots, T,$$

where $y^*_{1,t}$ is the latent monthly variables underlying the quarterly variables $y_{1,t}$ and Φ_j, $j = 1, \ldots, p$ is the (monthly - high frequency) coefficient matrix for the jth lag. Once identified the latent high frequency variables underlying the observed quarterly variables $y_{1,t}$, the MF-VAR is a standard VAR process for the monthly frequency, with classical specification of the coefficient matrices and with error covariance matrix Σ. As described in the next sections, this specification requires filtering procedures, in particular a state space representation and the use of the Kalman filter. The crucial point in MF-VAR is represented by the technical treatment of the unobservable values of the latent monthly series underlying the quarterly data. Different solutions have been provided.
In the following discussion we investigate different specifications of MF-VARs. We identify the alternative specifications presented by the literature, different estimation procedures and the developments provided in fifteen years of research.

1.2.1 Mariano and Murasawa's specification

The idea of Mariano and Murasawa (2003, 2010) is modelling MF-VARs by referring to a state space representation. Considering monthly and quarterly variables, the authors are interested in the analyses of the monthly dynamics between the observed monthly variables and the latent monthly processes underlying the quarterly variables.
Let $y_{1,t}, t = 1, \ldots, T$, be the N_1-variate processes of quarterly (low frequency) variables (observed only in the third month of the reference quarter), and let $y_{2,t}, t = 1, \ldots, T$, be the N_2-variate monthly (high frequency) process, sampled $m = 3$ times during each reference quarter. Consider the vector $y_t = (y'_{1,t}, y'_{2,t})'$, of dimension $N_1 + N_2 = N$. The MF-VAR investigates the dynamics between (i) the highest frequency (observed) variables $y_{2,t}$ and (ii) the high frequency latent processes $y^*_{1,t}$ underlying the observed low frequency series $y_{1,t}$. The specification of the model can be decomposed in two part: first, we have specify the autoregressive nature of the high frequency, not completely observable, autoregressive

process, and second, we have to define the relationship between the low frequency variables and the related unobservable high frequency counterpart.
Starting from this second aspect of the specification, the relationship between the vectors

$$y_t = \begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} \qquad \text{and} \qquad y_t^* = \begin{pmatrix} y_{1,t}^* \\ y_{2,t} \end{pmatrix},$$

can be generally written as

$$y_t - \mu = H(L)(y_t^* - \mu^*),$$

where L is the lag operator, $\mathbb{E}(y_t) = \mu$, $\mathbb{E}(y_t^*) = \mu^*$ and $H(L)$ is a polynomial of order k of weight matrices that link the quarterly observation of the subvector $y_{1,t}$ to k realizations of the monthly unobserved variables $y_{1,t}^*$. In the literature, the classical choices for temporal aggregation schemes move between point-in-time sampling and average sampling. The point-in-time sampling is considered for stock variables, for which the low frequency (quarterly) observation is obtained sampling the mth observation of the high frequency (monthly) process, for each reference low frequency period. On the other hand, considering $y_{1,t}$ as a flow variable we refer to the average sampling. In this case, the realizations of the quarterly variable are obtained as the average of the three monthly values of the underlying latent process, related to the reference quarter. Each temporal aggregation scheme determines the structure of the polynomial $H(L)$.
For sake of simplicity, assume that $\mu = \mu^* = 0$, then

$$y_t = H(L)y_t^*. \tag{1.1}$$

Following the scheme of decomposition of the specification, now we have specify the autoregressive nature of the high frequency, not completely observable, autoregressive process y_t^*. Assume that y_t^* follows a VAR(p) process, defined by

$$\begin{aligned} y_t^* &= \Phi_1 y_{t-1}^* + \Phi_2 y_{t-2}^* + \cdots + \Phi_p y_{t-p}^* + w_t \qquad (1.2) \\ \Phi(L)y_t^* &= w_t, \end{aligned}$$

with $w_t \sim WN(0, \Sigma)$, and $\Phi(L) = (I - \Phi_1 L - \Phi_2 L^2 - \cdots - \Phi_p L^p)$.
The MF-VAR is obtained considering a state space system, in which the measurement equation quantifies the relationship between the observable vector y_t and the underlying monthly process y_t^*, as defined in Eq. (1.1), and the specification of the state equation refers to the VAR process in Eq. (1.2). The state space model is given by the equations

$$y_t = Cs_t \tag{1.3}$$

$$s_{t+1} = As_t + B\epsilon_t, \tag{1.4}$$

with $w_t = B\epsilon_t$ and where Eq. (1.3) is the measurement equation and Eq. (1.4) is the state equation with $\epsilon_t \sim (0, I)$.
To define the matrix elements of the state space representation, we first pay attention to the measurement equation. The relationship quantified in Eq. (1.1) is rewritten in the form of Eq. (1.3) considering the state vector s_t and the input matrix C given by

$$s_t = \begin{pmatrix} y_t^* - \mu^* \\ y_{t-1}^* - \mu^* \\ \vdots \\ y_{t-k-1}^* - \mu^* \end{pmatrix} = \underbrace{\begin{pmatrix} y_t^* \\ y_{t-1}^* \\ \vdots \\ y_{t-k-1}^* \end{pmatrix}}_{N(k+1)\times 1}, \qquad C = \underbrace{\begin{pmatrix} H_0 & H_1 & \dots & H_k \end{pmatrix}}_{N\times N(k+1)} \tag{1.5}$$

where k is the order of the polynomial $H(L)$, and H_i, $i = 0, \dots, k$, is the weighting matrix related to the variable y_{t-i}^* in the definition of y_t (see Eq. (1.1)).
As mentioned above, the VAR(p) dynamic for y_t^* is considered in the state equation. In particular, we can rewrite the VAR(p) in Eq. (1.2) in the companion form as $s_t = As_{t-1} + B\epsilon_t$, with

$$s_t = \underbrace{\begin{pmatrix} y_t^* \\ y_{t-1}^* \\ \vdots \\ y_{t-p+1}^* \end{pmatrix}}_{Np\times 1}, \quad A = \underbrace{\begin{pmatrix} \Phi_1 & \Phi_2 & \cdots & \Phi_p \\ I_N & 0_{N\times N} & \cdots & 0_{N\times N} \\ 0_{N\times N} & \ddots & \cdots & \vdots \\ \vdots & \cdots & I_N & 0_{N\times N} \end{pmatrix}}_{Np\times Np}, \quad B = \underbrace{\begin{pmatrix} \Sigma^{\frac{1}{2}} \\ 0_{N\times N} \\ \vdots \\ 0_{N\times N} \end{pmatrix}}_{Np\times N}. \tag{1.6}$$

and the disturbances of the state equation $\epsilon_t \sim N(0, I_N)$.
As we can notice from the definition of the dimensions of the state vector in Eq. (1.5) and Eq. (1.6), the specification of the matrices in the state space system, and in particular, of the state vector s_t, is related to two non-negligible aspects: (i) how many lags of the latent process (in s_t) are involved in the analytical specification of the quarterly observation of $y_{1,t}$ (from the measurement equation), and (ii) the order of lags of the monthly VAR specified for s_t (from the transition equation). As illustrative case, assume that the order of the VAR is $p = 2$ and that the quarterly observations are given by the average sampling, i.e.

$$y_{1,t} = \frac{1}{3}(y_{1,t}^* + y_{1,t-1}^* + y_{1,t-2}^*).$$

Analytically, $y_{1,t}$ involves $k+1=3$ realizations of $y^*_{1,t}$, i.e. from Eq. (1.5) s_t is of dimension $3N \times 1$ with the polynomial $H(L)$ given by[6]:

$$H(L) = \begin{bmatrix} \frac{1}{3}I_{N_1} & 0 \\ 0 & I_{N_2} \end{bmatrix} + \begin{bmatrix} \frac{1}{3}I_{N_1} & 0 \\ 0 & 0 \end{bmatrix} L + \begin{bmatrix} \frac{1}{3}I_{N_1} & 0 \\ 0 & 0 \end{bmatrix} L^2,$$

whit N_1 is the number of low frequency variables, N_2 is the number of high frequency series, $N = N_1 + N_2$, and I_{N_i} with $i = 1, 2$ is the $N_i \times N_i$ identity matrix. Considering

$$\begin{aligned} C &= \begin{pmatrix} H_0 & H_1 & H_2 \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{3}I_{N_1} & 0_{N_1\times N_2} & \vdots & \frac{1}{3}I_{N_1} & 0_{N_1\times N_2} & \vdots & \frac{1}{3}I_{N_1} & 0_{N_1\times N_2} \\ 0_{N_2\times N_1} & I_{N_2} & \vdots & 0_{N_2\times N_1} & 0_{N_2\times N_2} & \vdots & 0_{N_2\times N_1} & 0_{N_2\times N_2} \end{pmatrix}, \end{aligned}$$

the measurement equation of the state space model is given by:

$$\begin{aligned} y_t &= C s_t \\ &= \begin{pmatrix} H_0 & H_1 & H_2 \end{pmatrix} \begin{pmatrix} y^*_t \\ y^*_{t-1} \\ y^*_{t-2} \end{pmatrix}, \end{aligned} \tag{1.7}$$

with $\mu^* = 0$, as above.
Considering Eq (1.6) and the order $p = 2$ of the VAR model, s_t is a $2N \times 1$ vector, with transition equation given by:

$$\begin{aligned} s_t &= A s_{t-1} + B\epsilon \\ \begin{pmatrix} y^*_t \\ y^*_{t-1} \end{pmatrix} &= \begin{pmatrix} \Phi_1 & \Phi_2 \\ I_N & 0_{N\times N} \end{pmatrix} \begin{pmatrix} y^*_{t-1} \\ y^*_{t-2} \end{pmatrix} + \begin{pmatrix} \Sigma^{\frac{1}{2}} \\ 0_{N\times N} \end{pmatrix} \epsilon_t. \end{aligned} \tag{1.8}$$

The dimensions of the state vector s_t in Eq. (1.7) and in Eq. (1.8) are different. In order to encompass the mismatch in the specification of the state vector, we identify two cases: the first for $p \le (k+1)$ and the second case for $p > (k+1)$.

CASE 1: p $\le$ (k+1). The number of elements in the state vector is defined by the order k of the polynomial $H(L)$, as in Eq. (1.5). In particular we consider

$$s_t = \underbrace{\begin{pmatrix} y^*_t - \mu^* \\ y^*_{t-1} - \mu^* \\ \vdots \\ y^*_{t-k} - \mu^* \end{pmatrix}}_{N(k+1)\times 1}$$

[6]See Mariano and Murasawa (2010) for the specification of the aggregation scheme in the relationship $y_t - \mu = H(L)(y^*_t - \mu^*)$, working with low frequency variables integrated of order 1.

with related matrices defined as:

$$A = \underbrace{\begin{pmatrix} \Phi_1 & \dots & \Phi_p & 0_{N\times N(k+1-p)} \\ & I_{Nk} & & 0_{Nk\times N} \end{pmatrix}}_{N(k+1)\times N(k+1)}, \tag{1.9}$$

$$B = \underbrace{\begin{pmatrix} \Sigma^{\frac{1}{2}} \\ 0_{Nk\times N} \end{pmatrix}}_{N(k+1)\times N}, \qquad C = \underbrace{\begin{pmatrix} H_0 & H_1 & \dots & H_k \end{pmatrix}}_{N\times N(k+1)}.$$

CASE 2: p > (k+1). The dimensions of the state vector s_t and (hence) of the matrices in the state space system, is determined by the lag order p of the VAR. Specifically,

$$s_t = \underbrace{\begin{pmatrix} y_t^* - \mu^* \\ y_{t-1}^* - \mu^* \\ \vdots \\ y_{t-p+1}^* - \mu^* \end{pmatrix}}_{Np\times 1},$$

$$A = \underbrace{\begin{pmatrix} \Phi_1 & \dots & \Phi_p \\ I_{N(p-1)} & & 0_{N(p-1)\times N} \end{pmatrix}}_{Np\times Np}, \tag{1.10}$$

$$B = \underbrace{\begin{pmatrix} \Sigma^{\frac{1}{2}} \\ 0_{N(p-1)\times N} \end{pmatrix}}_{Np\times N}, \qquad C = \underbrace{\begin{pmatrix} H_0 & \dots & H_k & 0_{N\times N(p-k+1)} \end{pmatrix}}_{N\times Np}.$$

At this point, we have introduced two components of the specification of a MF-VAR: (a.) an autoregressive behaviour for the unobservable monthly variables, and (b.) a relationship between the observed and the unobserved vectors. The next step refers to the treatment of the unobserved value of the low frequency series: if we consider that the quarterly variables are observed in the third months of each quarter, how can we treat the values of $y_{1,t}$ related to the first and second month of the quarters in the measurement vector? The general idea is to refer to these values as missing observations. To do this, we have to define a specific (*periodical time-varying*) structure. Hence, we specify the measurement equation such that:

$$y_{1,t}^{+} = \begin{cases} y_{1,t} & \text{if } y_{1,t} \text{ is observable} \\ \nu_t & \text{otherwise} \end{cases}.$$

with $\{\nu_t\} \sim \mathcal{N}(0, I_{N_1})$ and all realizations equal to zero. The result is a measurement vector rewritten as if the missing observations of $y_{1,t}$ were drawn from a standard Normal, independent of the model parameters. Then, in the general form, the (periodical time-varying) state space model is given by:

$$\begin{aligned} y_t^+ &= \mu_t + C_t s_t + D_t \nu_t \\ s_t &= A s_{t-1} + B\epsilon_t \end{aligned}$$

where, for all $t = 1, \dots, T$,

$$\begin{pmatrix} y_{1,t}^+ \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} \mu_{1,t} \\ \mu_2 \end{pmatrix} + \begin{pmatrix} C_{1,t} \\ C_2 \end{pmatrix} s_t + \begin{pmatrix} D_{1,t} \\ 0 \end{pmatrix} \nu_t,$$

$$\epsilon_t \sim N(0, I_N), \qquad \nu_t \sim N(0, I_{N_1}).$$

The (periodical time-varying) structure is defined by

$$\mu_{1,t} = \begin{cases} \mu_1 & \text{if } y_{1,t} \text{ is observable} \\ 0 & \text{otherwise} \end{cases}, \tag{1.11a}$$

$$C_{1,t} = \begin{cases} C_1 & \text{if } y_{1,t} \text{ is observable} \\ 0 & \text{otherwise} \end{cases}, \tag{1.11b}$$

$$D_{1,t} = \begin{cases} 0 & \text{if } y_{1,t} \text{ is observable} \\ I_{N_1} & \text{otherwise} \end{cases}. \tag{1.11c}$$

For the specified state space representation there's no further complication in applying Kalman filter and smoother, i.e. the (periodical time-varying) structure allows the filter to skip the missing observation. In particular, let F and G be two selection matrices, with dimensions defined by the two cases, identified in Eq. (1.9) and Eq. (1.10)

$$\begin{cases} F = \begin{pmatrix} I_N & 0_{N\times Nk} \end{pmatrix}, \quad G = \begin{pmatrix} I_{pN} & 0_{Np\times(k-p)N} \end{pmatrix}, & \text{if } p \leq k \text{ (CASE 1)} \\ F = \begin{pmatrix} I_N & 0_{N\times(p-1)N} \end{pmatrix}, \quad G = \begin{pmatrix} I_{Np} \end{pmatrix}, & \text{if } p > k \text{ (CASE 2).} \end{cases}$$

With the introduction of F and G, the VAR(p) process y_t^*, i.e. $\Phi(L)(y_t^* - \mu^*) = w_t$, with $w_t \sim N(0, \Sigma)$, can be rewritten as

$$\begin{aligned} F s_t &= y_t^* \\ &= \Phi_1 y_{t-1}^* + \cdots + \Phi_p y_{t-p}^* + w_t \\ &= \Phi G s_{t-1} + w_t \\ &= (I_N \otimes s_{t-1}' G') vec(\Phi) + w_t \\ F s_t - \Phi G s_{t-1} &= w_t, \end{aligned}$$

where $\Phi = (\Phi_1, \Phi_2, \dots, \Phi_p)$. Then, given the joint probability density function of the measurement and the state variables, i.e. $Y_t^* = (y_1^*, \dots, y_t^*)$ and $S_t = (s_0, \dots, s_t)$, with $Y_0^* = 0$, the log-likelihood function of the parameter vector $\theta = (vec(\Phi)', vech(\Sigma)')'$ is defined by

$$l(\theta; Y_T^*, S_T) = -\frac{NT}{2} ln 2\pi - \frac{T}{2} ln|\Sigma| - \frac{1}{2}\sum_{t=1}^{T}(Fs_t - \Phi G s_{t-1})'\Sigma^{-1}(Fs_t - \Phi G s_{t-1}).$$

The first step for obtaining EM estimates is running the Kalman filter. In particular, for $s_0 = 0$ (hence, $s_{1|0} = 0$ and $P_{1|0} = BB'$), and $t = 0, 1, \dots, T$, we run the updating and the prediction steps, respectively given by

$$\textbf{\textit{updating steps}} \qquad \begin{aligned} s_{t|t} &= s_{t|t-1} + K_t e_t \\ P_{t|t} &= (I_N - K_t H_t) P_{t|t-1}, \end{aligned}$$

where $K_t = P_{t|t-1}H'(HP_{t|t-1}H' + DD')^{-1}$ is the Kalman gain, and $e_t = y_t - As_{t|t-1}$.

$$\textbf{\textit{forecasting steps}} \qquad \begin{aligned} s_{t+1|t} &= As_{t|t} \\ P_{t+1|t} &= AP_{t|t}A' + BB'. \end{aligned}$$

Once obtained $s_{t|t-1}$ and $P_{t|t-1}$, we choose a starting value for the parameter vector θ, i.e. $\theta_0 = (vec(\Phi_0)', vech(\Sigma_0)')'$. The expectation step of the EM algorithm is obtained (with $y_0^+ = 0$) calculating, for $t = 1, \dots T$, the smoothed estimates

$$\begin{aligned} s_{t|T} &= \mathbb{E}(s_t|(y_1^+, \dots y_t^+)) \\ &= s_{t|t-1} + P_{t|t-1} r_t \\ P_{t|T} &= Var(s_t|(y_1^+, \dots y_t^+)) \\ &= P_{t|t-1} - P_{t|t-1} R_t P_{t|t-1} \\ P_{t+1,t|T} &= Cov(s_{t+1}, s_t|(y_1^+, \dots y_t^+)) \\ &= (I_N - P_{t+1|t} R_{t+1})(A(I_N - K_t H) P_{t|t-1} \end{aligned}$$

where, for $t = T, T-1, \dots, 1$,

$$\begin{aligned} r_t =& H'(HP_{t|t-1}H' + DD')^{-1}e_t + (A(I - K_t H))' r_{t+1}), \\ R_t =& H'(HP_{t|t-1}H' + DD')^{-1}H + (A(I_N - K_t H))' R_{t+1}(A(I_N - K_t H))), \end{aligned}$$

with $r_{T+1} = 0$, $R_{T+1} = 0$.
The maximization step consists in taking the conditional expectation of the score functions of the log-likelihood. In particular, from

$$\frac{\partial l(\theta; Y_T^*, S_T)}{\partial vec(\Phi)'} = \Sigma^{-1}\sum_{t=1}^{T} vec(Gs_{t-1}s_t'F' - Gs_{t-1}s_{t-1}'G'\Phi') = 0,$$

we calculate

$$\mathbb{E}\left\{\Sigma^{-1}\sum_{j=1}^{T} vec(Gs_{t-1}s_t'F' - Gs_{t-1}s_{t-1}'G'\Phi')\right\} = 0. \tag{1.13}$$

In order to solve the expectation in Eq. (1.13), we assume that, for all $k, j = 0, 1$,

$$M_{t-k,t-j} = \left\{\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}(s_{t-k}, s_{t-j}'|(y_1^+, \dots y_t^+))\right\}.$$

Then, the conditional expectation can be rewritten as

$$GM_{t,t-1}F' - GM_{t,t}G'\Phi' = 0,$$

obtaining the estimated coefficient matrix $\hat{\Phi} = (\hat{\Phi}_1, \dots \hat{\Phi}_p)$

$$\hat{\Phi} = (GM_{t,t}G')^{-1} GM_{t,t-1}F'.$$

Likewise, consider the score function obtained *w.r.t.* Σ^{-1}, and given by

$$\begin{aligned} &\frac{\partial l(\theta; Y_T^*, S_T)}{\partial \Sigma^{-1}} = \\ &= \frac{T}{2}\Sigma - \frac{1}{2}\sum_{t=1}^{T}(Fs_t s_t' F' - Fs_t s_{t-1}' G'\Phi' - \Phi G s_{t-1} s_t' F' + \Phi G s_{t-1} s_{t-1}' G'\Phi') = 0. \end{aligned} \tag{1.14}$$

Taking the expectation for the score function in Eq. (1.14), we obtain

$$\Sigma - (FM_{t-1,t-1}F' + FM_{t-1,t}G'\Phi' - \Phi GM_{t,t-1}F' + \Phi GM_{t,t}G'\Phi') = 0,$$

hence

$$\hat{\Sigma} = FM_{t-1,t-1}F' - FM_{t-1,t}G'(GM_{t,t}G')^{-1}GM_{t,t-1}F'.$$

After some iterations of the EM algorithm, the idea is to switch to a quasi-Newton method, i.e. use $\hat{\theta}_{EM} = (vec(\hat{\Phi})', vech(\hat{\Sigma})')'$ as initial value for the quasi-Newton estimation. Mariano and Murasawa motivate this choice, pointing out that "[...] *the EM algorithm slows down significantly near the maximum*".

1.2.2 Alternative Approaches

In this section we present the main alternative approaches proposed in the literature. The first three procedures consider state space representation. In particular, the crucial point (and the main difference between the three alternatives) is represented by the treatment of the missing data which arise considering mixed frequencies. The fourth alternative doesn't consider state space representation, but an extension of the Yule-Walker estimation approach.

Giannone, Reichlin and Small (2008)

In the framework of full-system methods, and in particular in the analyses of factor models, Giannone, Reichlin and Small (2008) focus on the evaluation of the effects of intra-monthly variables on the nowcasts of the GDP. Their idea is to update the nowcast of the GDP, each time a new information becomes available. The specification of Giannone, Reichlin and Small (2008) involves three crucial aspects of the nowcasting framework: the use of a large dataset of variables, which are available at high frequencies but with different time-delays in the publication, the requirement of a procedure that (i) summarizes the information of the initial dataset and (ii) maps the dynamics within the extracted factors and between the factors and the GDP.
The state space model considered by the authors extracts the common factors (obtained from a large dataset) in the measurement equation, and specifies a vector autoregressive process for the factors in the transition equation, i.e.

$$
\begin{aligned}
y_t &= \mu + \Lambda F_t + \xi_t, \qquad && WN(0, \Psi) \\
F_t &= AF_{t-1} + Bu_t, \qquad && u_t \sim WN(0, I_q)
\end{aligned}
$$

where y_t is a $N \times 1$ vector of quarterly variables obtained from monthly series through point-in-time sampling, $F_t = (f_{1,t}, \dots, f_{r,t})'$ is a $r \times 1$ vector of the r extracted factors from y_t, $\xi_t = (\xi_{1,t}, \dots, \xi_{1,N})'$ is the vector of idiosyncratic components, Λ is the $N \times r$ factor loading matrix and Ψ is a $N \times N$ diagonal matrix with entries ψ_i, $i = 1, \dots, N$. The main difference with Mariano and Murasawa (2003, 2010)'s strategy refers to the treatment of the missing observation: we don't need to replace missing values with zeros, but we set the variance of the measurement error to infinity. With a simple example, when a new information for the jth series of the initial dataset is considered, the other variables present missing observations. In this case the variances ψ_i, with $i = 1, \dots, N$, $i \neq j$ of the measurement error (idiosyncratic component of the factor model, summarized in the measurement equation) are set to infinity (missing observations). By this way, the Kalman filter provides the same result of Mariano and Murasawa (2003, 2010)'s procedure, i.e. the filter skips the missing observations.

The main advantage provided by this procedure regards the possibility of update the nowcast of the GDP each time a new observation (for each variable in the dataset) becomes available. From the estimated state vector $\hat{F}_t$, the nowcast of the GDP can be easily computed with the linear regression model given by $G\hat{D}P_t = \alpha + \beta'\hat{F}_t$, with α and β obtained from OLS estimation.

Schorfheide and Song (2015)

Durbin and Koopman (2012) consider the general problem of state space models with missing observations. The authors link the dimension of the observed vector to time, treating such a vector as a function of t. Schorfheide and Song (2015) refer to this last approach and deal with a mixed frequency vector of endogenous variables, with time-varying dimension.
Consider a simple case and specify a measurement vector containing variables sampled at two frequencies: N_1 quarterly series, i.e. $y_{1,t}$, appearing in the third month of the quarter, and N_2 monthly series, i.e. $y_{2,t}$, as described above. When the quarterly and the monthly instant coincide, the measurement vector is of dimension $(N_2 + N_1) \times 1$, i.e $y_t = (y'_{2,t} \quad y'_{1,t})'$, otherwise the procedure requires dealing with a $N_2 \times 1$ vector of monthly variables[7], i.e. $y_t = (y_{2,t})$
As in the previous sections, assume that y^*_t is the monthly VAR(p) process at which the economy evolves. Consider

$$s_t = \begin{pmatrix} y^*_t \\ y^*_{t-1} \\ \vdots \\ y^*_{t-p+1} \end{pmatrix}, \tag{1.15}$$

where y^*_t, $t = 1, \dots, T$, is driven by the VAR(p) process

$$y^*_t = \Phi_1 y^*_{t-1} + \cdots + \Phi_p y^*_{t-p} + w_t, \qquad w_t \sim iidN(0, \Sigma).$$

Assuming $\Phi = (\Phi_1, \dots, \Phi_p)'$, the related companion form of the model is given by

$$s_t = A_1(\Phi)s_{t-1} + \epsilon_t, \qquad \epsilon \sim iidN(0, \Omega(\Sigma)). \tag{1.16}$$

The quarterly variables in y_t can be rewritten as a "[...] *three-month average* [...] *only observed for every third month*", i.e.

$$\begin{aligned} \tilde{y}_{1,t} &= \frac{1}{3}(y^*_{1,t} + y^*_{1,t-1} + y^*_{1,t-2}) \\ &= \Lambda_1 s_t, \end{aligned}$$

[7]In accordance with Schorfheide and Song (2015), in this section we consider a state vector with the first elements represented by the observed monthly variables.

where Λ_1 is a block of the weighting matrix $\Lambda = (\Lambda_2, \Lambda_1)'$, with dimensions of the blocks related to N_2 and N_1, respectively, and $\tilde{y}_{1,t}$ is a monthly vector of observed and missing observations: the observed values refer to the third month of each quarter, and the missing observations are obtained in every t corresponding to the first and second months of the quarters.
The time-varying structure of $y_t = (y'_{2,t}, y'_{1,t})'$ is achieved introducing a sequence M_t of selection matrices. Identify $M_{1,t}$ as the matrices, in the sequence M_t, that are related to the quarterly variables in s_t. In the first and second month of each quarter $M_{1,t}$ is empty; therefore, at the third month, $M_{1,t}$ equals the identity matrix. By this way, $M_{1,t}$ allows to work with

$$\dim(y_{1,t}) = \begin{cases} N_1 & \text{if } \tilde{y}_{1,t} \text{ is observable} \\ 0 & \text{otherwise} \end{cases}$$

and $y_{1,t} = M_{1,t}\tilde{y}_{1,t} = M_{1,t}\Lambda_1 s_t$. Likewise, we can consider to include in M_t furthers selection matrices, in order to model different situations, as the time delay in publication, or the inclusion of additional monthly variables to the starting dataset[8]. In general, the result is a vector y_t with time-varying dimension, that allows to handle with the following measurement equation:

$$\begin{aligned} y_t &= M_t \Lambda s_t \\ &= \underbrace{\begin{pmatrix} M_{2,t} & 0 \\ 0 & M_{1,t} \end{pmatrix}}_{N\times N} \underbrace{\begin{pmatrix} \Lambda_2 \\ \Lambda_1 \end{pmatrix}}_{N\times N(p-1)} \underbrace{\begin{pmatrix} y_t^* \\ y_{t-1}^* \\ \vdots \\ y_{t-p+1}^* \end{pmatrix}}_{N(p-1)\times 1} \end{aligned}$$

with $t = 1, \ldots, T$, and transition equation defined in Eq. (1.16). The dimension of the weighting matrix Λ and the vector s_t are related to the VAR order p, as in Section 1.2.1. The state space described above is considered for $p \geq (k+1)$.
From the specified state space representation, Schorfheide and Song (2015) provide a Bayesian estimation procedure. Furthermore, in their empirical nowcasting problem, the authors highlight the improvement in the accuracy of the performances of short-horizons forecasts *w.r.t.* the results obtained with classical benchmark models (standard quarterly VAR) and with specific nowcasting tools, as MIDAS-type regressions.

[8]The procedure of the authors allows to include a wide range of situations, arising when the econometrician considers to deal with a real flow of mixed frequency variables. For instance, M_t could include a matrix $M_{j,t}$ for the introduction of further monthly variables, after a specific period T_b. The matrix $M_{j,t}$ will be empty for $t = 1, \ldots, T_b$, and equal the identity matrix for $t = T_b + 1, \ldots, T$.

Schorfheide and Song (2015) consider a Bayesian estimation procedure with proper Minnesota-style priors, for the treatment of their state space representation. The priors chosen by the authors are referable to the family of Multi-Normal Inverse-Wishart distributions[9]. Schorfheide and Song (2015) describe the conditional posterior densities of the VAR parameters and the state variables, these last introduced for achieving the mixed frequency characterization of the VAR. The choice of proper priors allows the authors to evaluate the posterior densities via a Gibbs sampling algorithm. For further details on Bayesian estimation of a state space model see Del Negro and Schorfheide (2011) and Giordani, Pitt and Kohn (2011).

Chen and Zadrozny (1998)

Another important contribution in the estimation of MF-VARs is furnished in Chen and Zadrozny (1998) and Koelbl, Braumanny, Felsensteinz and Deistler (2015). The authors consider extended Yule Walker (XYW) and maximum likelihood estimators of the parameters of a high frequency VAR model, obtained from mixed-frequency data.
Consider the vector $y_t = (y'^*_{1,t}, y'_{2,t})'$, of dimension $N_1 + N_2 = N$, where $y^*_{1,t}$ is the latent monthly variables underlying the quarterly variables $y_{1,t}$ and $y_{2,t}$ is the subvector of (completely observable) N_2 high frequency variables. We define the high frequency VAR(p) process as

$$y^*_t = \Phi_1 y^*_{t-1} + \Phi_2 y^*_{t-2} + \cdots + \Phi_p y^*_{t-p} + \varepsilon_t, \quad \varepsilon_t \sim iid(0, \Sigma), \quad t = 1, \ldots, T,$$

where Φ_j, $j = 1, \ldots, p$ is the (monthly - high frequency) coefficient matrix for the jth lag. Under regularity conditions, the (population) extended Yule Walker equations is given by

$$\begin{aligned} \mathbb{E}\left[(y_t)(y'_{2,t-1} \cdots y'_{2,t-1})\right] &= (\Phi_1, \ldots, \Phi_p)\mathbb{E}\left[\begin{pmatrix} y_{t-1} \\ \vdots \\ y_{t-p} \end{pmatrix} (y'_{2,t-1} \cdots y'_{2,t-1})\right] \\ Z_1 &= \mathbf{\Phi} Z_0. \end{aligned}$$

The solution

$$\begin{aligned} (\Phi_1, \ldots, \hat{\Phi}_p) &= Z_1 Z_0^{\dagger} \\ &= Z_1 Z'_0 (Z_0 Z'_0)^{-1} \end{aligned}$$

[9]For further details about Minnesota priors see Del Negro and Schorfheide (2011).

is obtained by replacing Z_1 and Z_0 with their sample estimates:

$$
\begin{aligned}
\hat{\gamma}_{2,2,T}(h) =& \mathbb{E}(y_{2,t+h}y'_{2,t}) = \frac{1}{T}\sum_{t=1}^{T-h} y_{2,t+h}y'_{2,t} \qquad h \geq 0 \\
\hat{\gamma}_{1,2,T}(h) =& \mathbb{E}(y_{1,t+h}y'_{2,t}) = \frac{1}{T/N}\sum_{t_1=1}^{t_2} y_{1,Nt}y'_{2,Nt-h}
\end{aligned}
$$

with $\hat{\gamma}_{2,2,T}(h) = \hat{\gamma}_{2,2,T}(-h)'$ and

$$
t_1 = \begin{cases} 1 & \text{if} \quad N > h \\ \lfloor \frac{h}{N} \rfloor & \text{if} \quad N \leq h \end{cases}, \qquad t_2 = \begin{cases} \lfloor \frac{T}{N} \rfloor & \text{if} \quad h \geq 0 \\ \lfloor \frac{T+h}{N} \rfloor & \text{if} \quad h < 0 \end{cases}.
$$

From the XYW estimator $\hat{\mathbf{\Phi}} = (\hat{\Phi}_1, \hat{\Phi}_2, \dots \hat{\Phi}_p) = \hat{Z}_1\hat{Z}_0^{\dagger}$, we can easily derive the GMM estimator

$$
vec(\hat{\mathbf{\Phi}}) = ((\hat{Z}_0 \otimes I_N)R_T(\hat{Z}'_0 \otimes I_N))^{-1}(\hat{Z}'_0 \otimes I_N)R_T vec(\hat{Z}_1).
$$

The relationship with the XYW estimator is given by $R_T = I_{N^2pN_2}$.
The estimator of the error covariance matrix can be obtained assuming

$$
G = \begin{pmatrix} I_N & 0 & \dots & 0 \end{pmatrix}
$$

and calculating

$$
vec(\Sigma) = \Big[(G \otimes G)(I_{(Np)^2} - (\hat{A} \otimes \hat{A}))^{-1}(G' \otimes G')\Big]^{-1} vec(\hat{\gamma}(0)),
$$

where

$$
\hat{A} = \begin{pmatrix} \hat{\Phi}_1 & \hat{\Phi}_2 & \cdots & \hat{\Phi}_p \\ I & 0 & \cdots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & \cdots & I & 0 \end{pmatrix}.
$$

Koelbl, Braumanny, Felsensteinz and Deistler (2015) consider both the cases of stock and flow low frequency variables, furnishing a generalization of the simplest case described above. They provide some properties of the XYW and GMM estimators, and identify the loss of information related to the analysis of "high frequency" and "mixed frequency" and between "mixed frequency" and "low frequency".

1.3 Mixed sampling frequency VAR models

Ghysels (2016) presents a new perspective for the analyses of mixed frequency data. The author considers the idea of specify the mixed frequency model referring to all the observations of the series during each reference low frequency period, without model the latent process underlying the observable realizations of the low frequency variables. We present below the main aspects of Mixed sampling frequency VAR models. Further details are presented in Ghysels (2016).

Consider the simple case of working with a quarterly variable ($N_1 = 1$), which appears at the end of the reference quarter, and a monthly variable ($N_2 = 1$) sampled $m = 3$ times during each low frequency period. In particular, we define τ the low frequency timely index, $y_{1,\tau}$ the low frequency variable, and $y_{2,\tau}$ the vector of the stacked m high frequency observations obtained during each low frequency period. Specifically we consider $y_{2,\tau} = (y_{2,\tau}^{(1)}, y_{2,\tau}^{(2)}, y_{2,\tau}^{(3)})'$[10]. For example, if we consider the first quarter of the year, the vector of high frequency observations will be given by

$$y_{2,\tau} = \begin{pmatrix} y_{2,\tau}^{(1)} \\ y_{2,\tau}^{(2)} \\ y_{2,\tau}^{(3)} \end{pmatrix} \quad \begin{matrix} January \\ February \\ March \end{matrix}$$

The general idea is to specify a low frequency VAR model for the endogenous vector y_τ obtained stacking (i) the m observations of the high frequency variables (available m times during the low frequency period), and (ii) the observation of the low frequency series. The resulting y_τ will be of dimension $N_S \times 1$, with $N_S = mN_2 + N_1$.

Coming back on the initial example, assuming that the observation of the quarterly variable appears only at the end of the reference quarter, the vector of endogenous variables will be given by the 4×1 vector defined by

$$y_\tau = \begin{pmatrix} y_{2,\tau} \\ y_{1,\tau} \end{pmatrix} = \begin{pmatrix} y_{2,\tau}^{(1)} \\ y_{2,\tau}^{(2)} \\ y_{2,\tau}^{(3)} \\ y_{1,\tau} \end{pmatrix}.$$

[10]The general specification of the high frequency sub-vector $y_{2,\tau}$, can be obtained considering

$$y_{2,\tau} = (y_{2,\tau}^{(1)}, \dots, y_{2,\tau}^{(k)}, \dots, y_{2,\tau}^{(m)})'$$

with apex $(j) = 1, \dots, m$, denoting the observation of the high frequency variable realized in the jth intra-quarter period: $j = 1$ if we refer to the first high frequency observation collected during the low frequency period, $j = 2$ if we refer to the second realization, and so on)

The *stacked skip-sampled* VAR(p) process, i.e. $y_\tau = A_0 + \sum_{j=1}^{p} A_j y_{\tau-1} + \epsilon_\tau$, can be specified as

$$\begin{pmatrix} y_{2,\tau}^{(1)} \\ y_{2,\tau}^{(2)} \\ \vdots \\ y_{2,\tau}^{(m)} \\ y_{1,\tau} \end{pmatrix} = A_0 + \sum_{j=1}^{p} A_j \begin{pmatrix} y_{2,\tau-j}^{(1)} \\ y_{2,\tau-j}^{(2)} \\ \vdots \\ y_{2,\tau-j}^{(m)} \\ y_{1,\tau-j} \end{pmatrix} + \epsilon_\tau, \tag{1.17}$$

where A_j, with $j = 1, \ldots, p$, are the $(mN_2 + N_1) \times (mN_2 + N_1)$ coefficient matrices of the VAR(p). By this way we don't refer to some latent high frequency process, underlying the low frequency variables, but we investigate the direct relationships between the endogenous variables of mixed frequency data at the lowest frequency. The generic matrix A_j of Eq. (1.17) can be explicitly written as

$$\begin{pmatrix} A_j^{1,1} & \cdots & A_j^{1,m} & A_j^{1,L} \\ \vdots & \cdots & \vdots & \vdots \\ A_j^{m,1} & \ldots & A_j^{m,m} & A_j^{m,L} \\ A_j^{L,1} & \ldots & A_j^{L,m} & A_j^{L,L} \end{pmatrix}$$

Ghysels (2016) discusses different ways to estimate the stacked MF-VAR, and quantifies the misspecification due to the estimation of a single (low) frequency VAR. The author formulates the maximum likelihood estimators of the parameters, the related spectral counterparts and a Bayesian estimation method, expanding the results obtained by Rodriguez and Puggioni (2010) and Ghysels and Owyang (2011), for Bayesian MIDAS-regression estimation[11].

Linear and non-Linear Systems

Similarly to the approach developed by Ghysels (2016), in the engineer literature of linear and non-linear systems some contributions to the analyses of mixed frequency data have been provided. Starting from the works of Bittanti (1986) and Bittanti and Colaneri (1991) about periodic systems, Chen, Anderson, Deistler and Filler (2012) investigate the relationships between blocked and unblocked representation of periodic linear systems, especially motivated by the use of the blocking techniques in generalized dynamic factor models with mixed frequency data.

[11] A similar approach has been provided by Blasques, Koopman and Malle (2015) for the specification of a dynamic factor model with mixed frequency data. In partial models, the counterparts of Ghysels (2016)' approach is represented by Carriero, Clark and Marcellino (2015), which propose a partial (Bayesian) model with stochastic volatility, with the aim of producing current quarter forecasts of GDP.

The blocking technique (as well known as lifting) has been applied in linear systems to transform linear discrete-time periodic systems to linear time-invariant systems. Consider the discrete (unblocked) system defined by

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t), \end{aligned}$$

where $t \in \mathbb{Z}$, $x(t) \in \mathbb{R}^n$, $y(t) \in \mathbb{R}^N$ and $u(t) \in \mathbb{R}^h$, with $N \geq h$. Let be $y(t) = (y_1(t)', y_2(t)')'$, where $y_2(t)$ is the N_2 variables sampled at the highest frequency, specifically m times during each low frequency period, and $y_1(t)$ are the N_1 low frequency variables (as above, $N_1 + N_2 = N$), appearing only at the end of the period. In blocked systems with mixed frequency data, we assume that the measurement vector is of dimension $mN_2 + N_1$ and is given by

$$Y(t) = \begin{pmatrix} y_2(t) \\ y_2(t+1) \\ \vdots \\ y_2(t+m-1) \\ y_1(t) \end{pmatrix}, \qquad t = 0, m, 2m, \dots T$$

with T multiple of m. Consider $C = (C_2', C_1')'$ and $D = (D_2', D_1')'$, where C_2 and D_2 are of dimensions $mN_2 \times n$ and $mN_2 \times mh$, respectively, and C_1 and D_1 are of dimensions $N_1 \times n$ and $N_1 \times hm$. The blocked system, provided to deal with mixed frequency data, is given by:

$$\begin{aligned} x(t+m) &= A_c x(t) + B_c U(t) \\ Y(t) &= C_c x(t) + D_c U(t), \end{aligned}$$

where $A_c = A^m$, $B_c = [A^{m-1}B, A^{m-2}B, \dots, B]$, $C_c = [C_2', A'C_2', \dots A'^{m-1}C_2', C_1']'$, $U(t) = (u(t)', u(t+1)', \dots, u(t+m-1)')$, and

$$D_c = \begin{pmatrix} D_2 & 0 & \cdots & 0 \\ C_2 B & D_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ C_2 A^{m-2} B & C_2 A^{m-3} B & \cdots & D_2 \\ D_1 & 0 & \cdots & 0 \end{pmatrix}.$$

The result is an explicit relationship between the blocked and unblocked systems, allowing to extend the knowledges about linear time-invariant systems to the analysis of linear discrete-time periodic systems.

Chen, Anderson, Deistler and Filler (2012) compare blocked and unblocked systems, analysing the classical properties of observability, controllability and minimality of linear systems. Furthermore, they consider the relationship between finite zeros[12] of the transfer functions of blocked and unblocked systems, highlighting some sufficient and necessary conditions for the presence of the zeros in blocked systems.

1.4 The Literature of MF-VAR models

In this section we present some contributions to the literature about the numerous developments of MF-VARs provided by researchers.

Markov-switching MF-VAR

Camacho (2013) introduces a Markov-switching dynamics in the MF-VAR model of Mariano and Murasawa (2003), in order to detect busyness cycle signals and probabilities of recession. The author allows the state space matrices to change over time, *w.r.t.* the (*periodical*) time-varying structure, specified by Mariano and Murasawa (2003) and reported in Eq. (1.11), and to evolve by a transition probability, determined by an irreducible 2-state Markov-switching chain.
The mixed frequency Markov Switching VAR (hereafter, MS-MF-VAR) can be represented by the equations

$$\begin{aligned} y_t &= C_{g_t} s_t + v_t \qquad v_t \sim iN(0, R), \\ s_t &= \mu_{g_t} + A_{g_t} s_{t-1} + B\varepsilon_t \qquad \varepsilon_t \sim iN(0, Q_{g_t}), \end{aligned}$$

with transition probability, associated to the two-state unobservable variable g_t, defined by

$$p(g_t = j/g_{t-1} = i) = p_{ij}.$$

The (*periodical*) structure specified to deal with missing observations in Eq. (1.11) for MF-VAR, is expanded in the MS-MF-VAR, and specifically regards y_t, C_{g_t}, v_t,

[12]The notion of zeros in linear system is related to the rank of the system matrix

$$M(Z) = \begin{pmatrix} ZI - A_c & -B_c \\ C_c & D_c \end{pmatrix},$$

with transfer function defined by $W(z) = C_c(ZI - A_c)^{-1}B_c + D_c$. The finite zeros of $W(z)$, with minimal realization of $\{A_c, B_c, C_c, D_c\}$, are the finite values of Z for which the rank of the system matrix $M(Z)$ decreases *w.r.t.* its normal rank. For further details about finite and infinite zeros see Grasselli and Longhi (1988) and Zamani, Chen, Anderson, Deistler and Filler (2011).

R and μ_{g_t}. The estimation of the transition probabilities can be obtained applying the nonlinear filter proposed by Hamilton (1989).

Foroni, Guerin and Marcellino (2014) extend the work of Camacho (2013) increasing the number of possible states and considering a Markov-switching structure for three classes of models: the MF-VAR specified by Mariano and Murasawa (2003), i.e MSMF-VAR (KF), the stacked MF-VAR of Ghysels (2012), i.e. MSMF-VAR (SV), and the MIDAS-type equations, i.e. MS-MIDAS. The authors emphasize the improvement in the accuracy of depicting turning points with MSMF-VAR (KF) models. They also consider a classical nowcasting exercise for: the models described above, the linear counterparts, i.e. MF-VAR (KF), MF-VAR (SV), MIDAS, Augmented Distributed Lags MIDAS (ADL-MIDAS), and for some benchmark models (as AR and VAR models). Furthermore, the authors combine the forecasts previously obtained for different macroeconomic indicators, finding that, in their specific exercise, the MSMF-VAR (KF) outperforms the other approaches, especially in terms of density forecasts.

MF-VAR with stochastic volatility

The inclusion of the stochastic volatility in a full system model is considered by Marcellino, Porqueddu and Venditti (2015). Specifically, they provide a dynamic factor model, with mixed frequency data, estimated with Bayesian methods. The proposed model is based on the specification of Mariano and Murasawa (2003).

The introduction of stochastic volatility corresponds to add a further equation to the starting state space system reported in Eq. (1.3) and Eq. (1.4). In particular, the new system is given by

$$\begin{aligned}
y_t =& Cs_t + D\varepsilon_t \\
s_{t+1} =& As_t + \epsilon_t, \qquad \epsilon_t \sim N(0, Q_t) \\
\epsilon_t =& R^{-1}\Lambda_t^{\frac{1}{2}} u_t, \qquad u_t \sim N(0, I) \\
vec(\Lambda_t) =& vec(\Lambda_{t-1}) + \zeta_t, \qquad \zeta_t \sim N(0, \Theta)
\end{aligned}$$

where Q_t is the (diagonal) error covariance matrix[13] defined by

$$Q_t = \mathbb{E}(R^{-1}\Lambda_t^{\frac{1}{2}} u_t u_t' \Lambda_t'^{\frac{1}{2}} R^{'-1}) = R^{-1}\Lambda_t R'^{-1}.$$

By this way, $\Lambda_t = RQ_tR'$ is a matrix of drifting volatilities, and Θ is a diagonal covariance matrix. The diagonal entries of Q_t are considered as functions of the

[13]The MF-VAR specified in Section 1.2 presents a time-invariant covariance matrix of the error of the measurement equation.

elements of Λ_t.

The Mariano and Murasawa (2003)'s specification can be considered as nested in the Marcellino, Porqueddu and Venditti (2015)'s model: the dynamic factor model is (intuitively) a MF-VAR in state space representation, with state variables given by the factors extracted from a large dataset. Marcellino, Porqueddu and Venditti (2015) consider a Bayesian estimation procedure, with informative priors on steady state means. In their forecasting exercises they find results which are consistent with the mixed frequency data literature: the use of a mixed frequency set-up provides some non-negligible improvements in the accuracy of the forecasts; significant enhancements are especially obtained considering short horizons.

Factor augmented MF-VAR

Marcellino and Sivec (2015) consider FAVAR models and structural FAVARs, with mixed frequency data. The work is presented as extensions of: (i) Mariano and Murasawa (2010)'s specification and (ii) Doz, Giannone and Reichlin (2006)'s estimation procedure.
The mixed frequency FAVAR model corresponds to the state space system

$$\begin{aligned} y_t =& C s_t + \nu_t \\ s_{t+1} =& A s_t + B \epsilon_t \end{aligned}$$

where, differently from Eq. (1.3) and Eq. (1.4), s_t contains the factors extracted from a large dataset, $C = (H\Lambda, 0, \dots, 0)$ contains the product of the temporal-aggregation matrix H and the factor loadings matrix Λ of dimension $N \times k$, where k equals the number of extracted factors and ν_t is a compound error term, i.e. $\nu_t = He_t + Dv_t$. In particular, ν_t is the sum of the temporal aggregation of the idiosyncratic components of the extracted factors, i.e. He_t, and a second component, i.e. Dv_t.
The most important source of indeterminacy is related to the extraction of the factors. This procedure, in fact is as well known affected by some identification problems. In the literature different restricting procedures have been applied. One of the most classical choice is considering a block of the factor loadings matrix equals the identity matrix.
In the MF-FAVAR a structural dynamic can be inserted considering some shocks in the observable factors. The idea is associating the shocks to those factors for which the sub-matrix of factor loadings is restricted.

Marcellino and Sivec (2015) compare the estimates and the impulse responses obtained with their specification with those obtained with previous empirical analysis of the VAR and FAVAR literature, i.e. Bernanke, Boivin and Eliasz (2005), Bernanke, Gertler, Watson, Sims and Friedman (1997) and Ramey (2011). The authors find some interesting differences with the existing results. Furthermore, Marcellino and Sivec (2015) study the effects of credit shocks on quarterly GDP in their specific application. They emphasize the ability of the MF-SFAVAR to overcome some limitations of more parsimonious methodologies, highlighting both differences and similarities *w.r.t.* the models presented in the literature.

Time-varying MF-VAR

In the literature of MF-VAR some contributions have been proposed to allow the model dynamics to vary over-time, with continuously time varying parameters. In particular, Cimadomo and D'Agostino (2015) consider a time-varying structure for the Mariano and Murasawa (2003)'s specification and Götz and Hauzenberger (2015) present a time-varying MF-VAR (hereafter, TV-MF-VAR) starting from the Schorfheide and Song (2015)'s model.
In general, a TV-MF-VAR model can be expressed by

$$\begin{aligned}
y_t =& C s_t + v_t, \qquad v_t \sim N(0, R_t) \\
s_{t+1} =& A_{0,t} + A_t s_t + \epsilon_t, \qquad \epsilon_t \sim N(0, \Sigma_t) \\
\Sigma_t =& F_t D_t F_t',
\end{aligned}$$

where $A_{0,t}$ is the vector of time-varying intercepts, R_t is a diagonal matrix, F_t is lower triangular, with ones on the main diagonal and D_t is a diagonal matrix. The motions of the time-varying components of the system are summarized by the following equations

$$\begin{aligned}
\theta_t =& \theta_{t-1} + \omega_t, \qquad \omega_t \sim N(0, \Omega) \\
\log \sigma_t =& \log \sigma_{t-1} + \zeta_t, \qquad \zeta_t \sim N(0, \Xi) \\
\phi_t =& \phi_{t-1} + \psi_t, \qquad \psi_t \sim N(0, \Psi),
\end{aligned}$$

where $\theta_t = (vec(A_{0t})'vec(A_t)')'$, σ_t is the vector of diagonal entries of $D_t^{1/2}$ and ϕ_t is the vector of the non-zeros off-diagonal elements of F_t^{-1}, related to the contemporaneous relationships among each equation.

Both Cimadomo and D'Agostino (2015) and Götz and Hauzenberger (2015), consider a Bayesian estimation procedure. The latter allow only for a time-varying structure in the intercepts and in the error variances; however they provide an approximate estimation procedure, suggested by the *online predictions* in engineering literature, in order to alleviate the computational effort due to non-small scale TV-MF-VAR.

VECM and cointegration for mixed frequency data

For the analysis of cointegrated MF-VAR and MF-Vector Error Correction Models, several papers have been proposed.
Starting from a N-variate monthly VAR(p) $\sim I(1)$, i.e. $y_t^* = \Phi_1 + y_{t-1}^* + \cdots + \Phi_p y_{t-p}^* + \varepsilon_t$, for $t = 1, \ldots, T$ and with cointegration rank r, $0 < r < N$, we consider the related Error Correction form given by

$$\begin{aligned} y_t^* - y_{t-1}^* = \triangle y_t^* =& \Pi^* y_{t-1}^* + \sum_{j=1}^{p-1} \Psi_j^* \triangle y_{t-j}^* + \varepsilon_t \\ =& \alpha^* \beta'^* y_{t-1}^* + \sum_{j=1}^{p-1} \Psi_j^* \triangle y_{t-j}^* + \varepsilon_t \end{aligned} \tag{1.19}$$

where $\Pi^* = -(\Phi) = -(I - \Phi_1 - \cdots - \Phi_p)$, $\Psi_i^* = \sum_{i=j+1}^{p} \Phi_i$ and $\varepsilon_t \sim N(0, \Omega)$. The identification of the model is guaranteed ordering y_t^* such that the no-cointegrated variables are all in the last $n - r$ position of the vector. This arrangement allows to consider the normalization $\beta^* = (I_r, \tilde{\beta}'^*)'$, with $\tilde{\beta}^*$ of dimension $(n - r) \times r$.
In order to consider a mixed frequency data framework in the specified VECM, we introduce the permutation matrix P, such that the vector of mixed frequency observed variables $y_t = P y_t^*$. Then multiplying both side of Eq. (1.19) by P, we obtain a MF-VECM defined by

$$\triangle y_t = \alpha \beta' y_{t-1} + \sum_{j=1}^{p-1} \Psi_j \triangle y_{t-j} + \epsilon_t \tag{1.20}$$

where $\alpha = P\alpha^*$, $\beta = \beta'^* P$, $\Pi = P\Pi^*$, $\Psi = P\Psi^*$ and $\epsilon_t = P\varepsilon_t \sim N(0, P\Omega P')$. Casting the solution in state space form, we obtain

$$\begin{aligned} y_t =& C s_t + v_t \\ s_{t+1} =& A s_t + B \epsilon_t \end{aligned}$$

where

$$s_t = \begin{pmatrix} y_t^* \\ y_{t-1}^* \\ \vdots \\ y_{t-p+1}^* \end{pmatrix}, A = \begin{pmatrix} (I + \alpha\beta' + \Psi_1) & (\Psi_2 - \Psi_1) & \cdots & (\Psi_{p-1} - \Psi_{p-2}) & -\Psi_p \\ I_n & 0 & \cdots & \cdots & 0 \\ 0 & I_n & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \cdots & I_n & 0 \end{pmatrix},$$

and $B = (I_n, 0)'$. The estimation of the MF-VECM with the EM algorithm is described in Seong, Ahn and Zadrozny (2013).

Seong, Ahn and Zadrozny (2013) provide a Monte Carlo analysis and an empirical application of their proposed estimation procedure via EM algorithm. Götz, Hecq and Ubrain (2013) consider common cycles in the MF-VAR version of Ghysels (2012), described below. The authors extend the stacked MF-VAR to work with I(1) variables, and consider an alternative model, in order to provide a Likelihood-Ratio-based test for the presence of common cyclical features. Ghysels and Miller (2015) consider the MF-VAR of Mariano and Murasawa (2003, 2010), and provide a test to detect cointegration, both with temporally aggregated and mixed frequency data.

1.5 Structural MF-VARs

A structural MF-VAR (hereafter, MF-SVAR) is a mixed frequency VAR, adapted for the analyses of structural dynamics. Starting from classical VAR representation, we can write an exemplifying trivariate MF-SVAR(1) model as

$$\begin{pmatrix} y_{1,t}^* \\ y_{21,t} \\ y_{22,t} \end{pmatrix} = \underbrace{\begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix}}_{A} \begin{pmatrix} y_{1,t-1}^* \\ y_{21,t-1} \\ y_{22,t-1} \end{pmatrix} + \underbrace{\begin{pmatrix} b_{11} & 0 & 0 \\ b_{12} & b_{22} & 0 \\ b_{13} & b_{23} & b_{33} \end{pmatrix}}_{B} \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{21,t} \\ \epsilon_{22,t} \end{pmatrix}, \tag{1.21}$$

where $y_{21,t}$ and $y_{22,t}$ are two monthly endogenous variables, $y_{1,t}^*$ is the latent monthly process underlying a quarterly variable $y_{1,t}$, $\epsilon_t = (\epsilon_{1,t}, \epsilon_{21,t}, \epsilon_{22,t})' \sim N(0, I_3)$ is the vector of structural shocks, and B is a lower triangular matrices, obtained from the Choleski decomposition of the error covariance matrix Σ_u of $u_t = B\epsilon_t$. Classical identification framework can be investigated considering some aspects of the problem. As mentioned before, we can assume different relationships between quarterly variables and the related underlying monthly latent process. For instance, consider $y_{1,t}$ as the result from a point-in-time sampling procedure of

$y^*_{1,t}$; assuming the same approach for the monthly variables, we can obtain the quarterly realization of the monthly series, obtained with point-in-time sampling. By this way, we can estimate a (completely observable) quarterly VAR process. This specification is analytically related to the unobservable monthly process of Eq. (1.21). In particular, starting from a generic monthly SVAR(1)

$$Y^*_t = AY^*_{t-1} + B\varepsilon_t, \qquad \varepsilon_t \sim (\mathbf{0}, I_3), \tag{1.22}$$

we introduce the polynomial $\Gamma(L) = (I + AL + A^2L^2)$. Multiplying both sides of Eq. (2.2) for the polynomial $\Gamma(L)$, we obtain the expression

$$\begin{aligned}(I + AL + A^2L^2)(I - AL)Y^*_t &= (I + AL + A^2L^2)B\varepsilon_t \\ (I - A^3L^3)Y^*_t &= (I + AL + A^2L^2)B\varepsilon_t \\ Y_\tau &= A^3Y_{\tau-1} + \xi_\tau \end{aligned} \tag{1.23}$$

with $\tau = 3t$ (*quarters*) and

$$\xi_\tau \sim (0, \Omega), \qquad \Omega = BB' + ABB'A' + A^2BB'A'^2.$$

The polynomial $\Gamma(L)$ ensures a quarterly solution in which the endogenous variables are related only to their third (observable) lags.
Eq. (1.23) represents the (completely observable quarterly) VAR process[14], generated with point-in-time sampling of the monthly variables in the VAR of Eq. (2.2).

Foroni and Marcellino (2014, 2016) treat structural MF-VAR and identification issues. The authors computationally investigate a SVAR(1) process, and explicit the consideration just proposed in Marcellino (1999), about some consequences of aggregation procedures. Foroni and Marcellino (2016) find that, considering a MF-VAR for one quarterly series and two monthly variables, we can obtain a full identified monthly specification of the model. Foroni and Marcellino (2014) proposed the same approach for a small scale DSGE model.
The specific issue of identification of structural MF-VAR was also addressed by Kim (2010). He investigates the problems rising from frequency misspecification. In particular, he refers to a data augmentation approach in a Bayesian framework for the treatment of mixed frequencies.

A different specification of the problem has been provided by Ghysels (2016), from the starting specification of mixed sampling frequency VARs presented in Section 1.3

[14]In the case of mixed frequency data, the standard approach consists of transform, through aggregation or point-in-time sampling, the high frequency series in a unique frequency (the lowest frequency).

1.6 Conclusions

In the last fifteen years, many contributions have been provided by the literature to deal with mixed frequency data. Banbura, Giannone, Modugno and Reichlin (2013) provide an exhaustive survey of the econometric tools furnished in the last decade. An important role in the analyses of the relationships between mixed-frequency data, is played by mixed-frequency autoregressive processes (MF-VAR). Considering a set of data collected at two frequencies, a MF-VAR is a standard (but, unobservable) VAR process for all the variables treated at the high frequency. In particular, for the low frequency series, we assume that exists an underlying (latent) high frequency process, which can be investigated with state space models and filtering procedures.
Motivated by different aspects of classical empirical problems, especially by the overwhelming framework of nowcasting, the researchers investigate different aspects of the MF-VAR models, as the inclusion of stochastic volatility, the analyses of Markov switching MF-VARs, cointegration and structural relationships. The result is an ever-expanded wide range of powerful, econometric tools, increasingly employees in the most varied empirical applications.

Chapter 2

A moment-based approach for identification and estimation of mixed frequency Structural VARs.

2.1 Introduction

In the analysis of the co-movements between economic series, working with variables sampled at different frequency is a common situation. The classical approach is represented by the estimation of the model at the lowest frequency in the data. In other words, the high frequency variables are aggregated until all series present the same low frequency. After that, the model is classically estimated (*naive approach*). However, as pointed out by Marcellino (1999)[1], temporal aggregation can lead to different problems of identification, estimation and interpretation of the results due to the misspecification of the co-movements between mixed frequency variables[2] (*aggregation bias*). In particular, the author sums up the characteristics of a time series process that appear to be either unaffected or affected by temporal aggregation. The structural analysis, and specifically the Impulse Response functions (IRFs) belong to the second category, i.e. the set of time series "properties" that vary after aggregation. For example, consider to work with two variables, one available at monthly frequency and a second one available only every three months (quarterly). We consider to estimate a structural VAR model (SVAR), with a recursive identification scheme, in order to analyse the response of the first variable

[1]See also Wei (1981), Weiss (1984), Rossana and Seater (1992), Pierse and Snell (1995), Kim (2010).

[2]See Ghysels (2016).

to an orthogonal shock in the second. With the classical approach, the econometrician aggregates the monthly variable, and estimates a quarterly SVAR, with related quarterly IRFs. Now assume that we can deal also with the monthly realizations of the quarterly variable. We estimate the monthly SVAR, and then the monthly IRFs. In order to compare the results obtained with the two approaches, we aggregate the monthly responses, and we plot both the quarterly solutions: the IRFs for the quarterly process (*naive approach*), reported in Figure 2.1 with the black solid line, and the aggregated high frequency impulse responses, shown with the red dotted line.
As shown in Figure 2.1 the responses are quite different. In particular, the magnitude of the instantaneous response obtained with the aggregated solution is higher then that obtained with the *naive* approach.

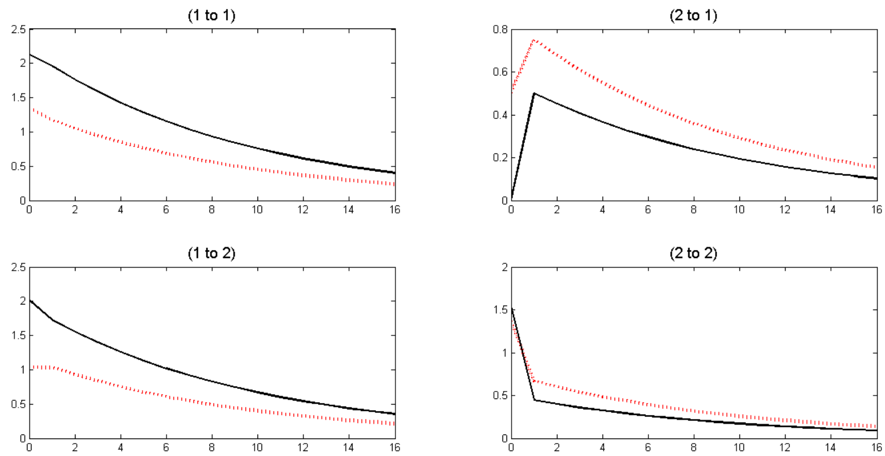


Figure 2.1: Comparison between low frequency Impulse Responses, calculated with the naive approach (black solid line), and the responses obtained aggregating the estimated high frequency IRFs (red dotted line).

In the last decade, the researchers try to implement different econometric solutions to use all the information in mixed frequency datasets and, therefore, to mitigate the effects of temporal aggregation. One of the proposed approaches considers the general idea of extend the VAR methodology to mixed frequency data. Mixed frequency VAR models (MF-VARs) have been introduced more than twenty years ago by Harvey and Pierse (1984) and Zadrozny (1988). The methodology has been finally refined and developed by Mariano and Murasawa (2003, 2010). A MF-VAR can be defined as a high frequency VAR process, through which we analyse the co-movements between the (available) high frequency data and the latent high frequency processes underlying the low frequency variables. As widely analysed in Chapter 1, the general idea of MF-VAR consists in referring to a VAR process in which the endogenous vector is not completely observable. State space representation handles both the latent and the autoregressive characteristic of the MF-VAR. In particular, we can write a MF-VAR process as

$$y_t^* = \Phi_1 y_{t-1}^* + \Phi_2 y_{t-2}^* + \cdots + \Phi_p y_{t-p}^* + \varepsilon_t, \quad \varepsilon_t \sim iid(0, \Sigma), \quad t = 1, \ldots, T, \quad (2.1)$$

where $y_t^* = (y_{1,t}^{*\prime}, y_{2,t}^{\prime})^{\prime}$ is the $N \times 1$ vector of high frequency endogenous variables: $y_{2,t}$ is the subvector of N_2 variables observed each high frequency instant, and $y_{1,t}^*$ is the subvector of N_1 high frequency latent variables underlying the quarterly variables $y_{1,t}$; Φ_j, $j = 1, \ldots, p$ is the (monthly - high frequency) coefficient matrix for the jth lag, and Σ the error covariance matrix. Reparametrizing Eq. (2.1) with state space representation, in the state equation we refer to the high frequency autoregressive nature of the process, while in the measurement equation we specify the relationships between the latent high frequency process (i.e. $y_{1,t}^*$) underlying the observable variables, and the only available low frequency realizations (i.e. $y_{1,t}$).
If in MF-VARs the frequency investigated by the researchers is the highest frequency in the data, a different approach is presented by Ghysels (2016). The author proposes a novel specification of VAR processes with mixed-data sampling. The idea of Ghysels is strongly inspired by the MIDAS literature[3]. MIDAS regressions aggregate data sampled at different frequencies in order to analyse the relationships between the low frequency variable (of interest) and some regressors sampled at different frequency. The aggregation is done through a parsimonious weighting function, which depend on a low dimensional parameter vector, and links the low frequency variable to the high frequency observations of the regressors. By this way, Ghysels (2016) consider the idea of modelling a VAR process in which the endogenous vector is obtained stacking both the high and the low frequency observations in the same vector, in order to study the impact of high/low frequency data on low/high frequency variables.

Among the developments provided about MF-VARs[4], only a small field of the literature considers the structural analysis. After Foroni and Marcellino (2014) for the estimation of dynamic stochastic general equilibrium models (DSGE) with mixed frequency data, a more general investigation of the MF-Structural VAR

[3]Mixed-Data Sampling equation (MIDAS) represents one of the latest tool provided in the framework of nowcasting; see Ghysels, Santa-Clara and Valkanov (2006), Ghysels, Sinko and Valkanov (2006), Clements and Galvão (2008)), among others.

[4]See section 1.4.

(MF-SVAR) is presented by Foroni and Marcellino (2016). The authors consider the simple cases of a VAR(1) models, and analyse the relationships from the high and the low frequency representation of the problem. They refer to two different aggregation schemes: the point-in-time sampling and the aggregation sampling. Foroni and Marcellino demonstrate that, considering a suitable estimation procedure, the analysis of mixed frequency data alleviates the classical bias due to temporal aggregation.
In this discussion we present a novel approach that doesn't belong completely to one of the classes of MF-VAR identified before, and represented by Mariano and Murasawa (2003, 2010)'s approach and Ghysels (2016)'s specification. The proposed approach allows to recover high frequency parameters from low frequency estimates, considering the mapping between the matrices. Specifically, starting with a general high frequency SVAR(1) process, aggregated in different ways, we demonstrate that we can recover the high frequency parameters, starting from the estimated low frequency counterpart, through a Minimum Distance estimation approach. Referring to a VAR of order one could be seems to involve only a small part of the multivariate time series literature. However, in the macroeconomic framework an important field of research, represented by Liner Expectation models and Dynamic Stochastic General Equilibrium (DSGE) models, usually refer to VAR(1) processes.
Motivated by this strand of the literature, in Section 2.2 we introduce the general specification of the problem, the aggregation schemes which we refer, and the generalization proposed in this discussion. Section 2.3 states the first results obtained from the point-in-time sampling technique. The estimation procedure is introduced, pointing out some critical point about identification and high nonlinearity of the relationships. We propose some Monte Carlo experiments, and a comparison with the state space methodology. The second aggregation scheme, i.e. the sum-over-the quarter sampling, is introduced in Section 2.4. As in Section 2.3, we propose two different Monte Carlo experiments, and a comparison with the results obtained with the state space approach. Section 2.5 asses a possible generalization of the procedure to higher order high frequency VARs. We introduce the Impulse Response Function Matching estimator. The general procedure is described, and we propose a further Monte Carlo experiment to evaluate the generalization approach. In Section 2.6 we report the results of some empirical exercises of structural analysis, with different specification of the problem. In Section 2.7 we provide some final considerations and conclusions.

2.2 Identification of the high frequency structural VAR(1)

In this section we introduce the general MF-SVAR(1), considered by Foroni and Marcellino (2014, 2016), and the estimation procedure suggested by Giannone, Monti and Reichlin (2014) to solve the identification problem. We focus on the following monthly trivariate ($n = 3$) structural VAR:

$$Y_t^* = AY_{t-1}^* + B\varepsilon_t, \qquad \varepsilon_t \sim (\mathbf{0}, I_3), \tag{2.2}$$

where $Y_t^* = (y_{1,t}'^*, y_{2,t}', y_{3,t}')'$, and $y_{1,t}^*$ is the variable observed only at quarterly frequency, B is a $n \times n$ matrix related to the $n \times 1$ vector of structural shocks ε_t, and $u_t = B\varepsilon_t$ is the residual component. In the discussion of Foroni and Marcellino (2016), B is obtained from a Cholesky decomposition of the covariance matrix Σ_u of the residual component. In particular, the matrix B governs the relationships between the disturbances and the structural shocks. The analysis can be generalized to SVARs with generic n and, in particular cases, to non-recursive structural identification schemes.
We introduce:

- the filter $\Gamma(L) = (I + AL + A^2L^2)$, such that, multiplying both sides of Eq. (2.2) by $\Gamma(L)$, we obtain only power of degree $m = 3$ of the lag operator L (with m equal to the number of high frequency observations obtained in each reference low frequency period), $L^3Y_t = Y_{t-3}$. By this way we refer only to observations corresponding to the third-monthly values available at each quarter;

- $\omega(L)$ the weighting matrix, that governs the sampling (aggregation) approach[5]. Throughout this paper, we consider two cases of interest:

$$\text{point-in-time sampling} \qquad \omega(L) = I,$$

$$\text{sum (-over-the low frequency) sampling} \qquad \omega(L) = (I + L + L^2).$$

The choice of the aggregation scheme is substantially connected to the nature of the variables. In particular, if we have to deal with stock variables we will refer to the *point-in-time* sampling; on the other hand, in the case of flow variables we aggregate the data with the *sum-over the low frequency* technique[6]. The two sampling procedures provide distinct solutions. For this reason, and to make our presentation as simple as possible, we divide the problem in two parts: in the first

[5]Examples of sampling procedure have been provided in Chapter 1.
[6]Directly linked to the "sum" approach, is the (most common) average sampling.

one we consider "point-in-time" sampling, and in the second the *"sum-over-the quarter"* case. For brevity, we refer to this technique as "sum" sampling.
After the choice of the aggregation scheme, the general idea of the proposed approach is to recover the parameters of the (unobservable) high frequency process, investigating the mapping between the available low frequency (quarterly) parameters and the high frequency (monthly) counterparts.

2.3 Point-in-time sampling

Assume that the monthly DGP is described by the SVAR(1) in Eq. (2.2). In the point-in-time case the quarterly values of $y^*_{1,t}$ correspond to the third-period observations of the monthly process, in each reference quarter.
In order to obtain the aggregated solution, we multiply both sides of Eq. (2.2) by the weighting matrix $\omega(L)$ and the filter $\Gamma(L)$[7]. The resultant low frequency process is given by

$$\begin{aligned}
\Gamma(L)(I-AL)\omega(L)Y^*_t &= \Gamma(L)B\omega(L)\varepsilon_t \\
(I+AL+A^2L^2)(I-AL)(I)Y^*_t &= (I+AL+A^2L^2)B(I)\varepsilon_t \\
(I-A^3L^3)Y^*_t &= (I+AL+A^2L^2)B\varepsilon_t \\
Y_t &= A^3Y_{t-3} + B\varepsilon_t + AB\varepsilon_{t-1} + A^2B\varepsilon_{t-2} \\
Y_\tau &= CY_{\tau-1} + \xi_\tau \qquad (2.3)
\end{aligned}$$

where $\tau = 3t$ and $\xi_\tau = B\varepsilon_t + AB\varepsilon_{t-1} + A^2B\varepsilon_{t-2}$. Furthermore, since $\varepsilon_t \sim (0, I)$, $\xi_\tau \sim (0, \Omega)$ with

$$\Omega = BB' + ABB'A' + A^2BB'A'^2. \qquad (2.4)$$

and ξ_τ is uncorrelated in τ (quarterly frequency), i.e.

$$Cov(\xi_\tau, \xi_{\tau-l}) = 0, \qquad l = \pm1, \pm2, \ldots.$$

In general, the researcher estimates the econometric model at the frequency at which all data are available, which in this case is quarterly. As mentioned above,

[7]We can obtain the same results by repeated substitutions, i.e.

$$\begin{aligned}
Y^*_t &= AY^*_{t-1} + B\varepsilon_t \\
Y^*_t &= A(AY^*_{t-2} + B\varepsilon_{t-1}) + B\varepsilon_t \\
Y^*_t &= A(A(AY^*_{t-3} + B\varepsilon_{t-2}) + B\varepsilon_{t-1}) + B\varepsilon_t \\
Y^*_t &= A^3Y^*_{t-3} + (I+AL+A^2L^2)B\varepsilon_t \\
Y_\tau &= A^3Y_{\tau-1} + \xi_\tau.
\end{aligned}$$

in a dataset composed by variables sampled at different frequency, the typical approach is represented by the aggregation of the high frequency variables, until all the data presents the same (low) frequency. This strategy corresponds to assume that the Data Generating Process belongs to the model at the lowest data frequency. On the other hand, if we assume that the DGP belongs to the highest frequency, the estimate of the aggregated model could lead to relevant estimation bias[8].

In this discussion we assume that the DGP is at monthly (highest) frequency. We propose a two step procedure to obtain the estimates of $\hat{A}$ and $\hat{B}$ of Eq. (2.2), given the estimated matrices $\hat{C}$ and $\hat{\Omega}$ in Eq. (2.3) and Eq. (2.4), based on quarterly data.

The presented case refers to a structural matrix B lower triangular. However, in some situations, triangular SVARs offer a misspecified representation of the reality. Castelnuovo and Surico (2010), and Bacchiocchi, Castelnuovo and Fanelli (2016) present different examples for which a Cholesky decomposition of the residual covariance matrix leads to non-consistent estimation of the parameters and to unreliable responses of the variables after a policy monetary shock. To generalize the procedure described above, we evaluate the aggregation of SVAR(1) processes also for non-recursive structural schemes and we exploit how to achieve identification in those cases in which we observe identification problems.

2.3.1 Estimation approach

Imagine that the DGP belongs to the model at highest frequency t (monthly), but we can deal only with variables at the lowest frequency τ (quarterly). Our objective is to identify the monthly parameters in A and B, with the knowledge of the quarterly parameter matrices C and Ω. The monthly model specification is given by:

$$Y_t^* = AY_{t-1}^* + B\varepsilon_t, \quad \varepsilon_t \sim (\mathbf{0}, I_3), \quad t = 1, \dots T$$

with quarterly correspondent process:

$$Y_\tau = CY_{\tau-1} + \xi_\tau, \quad \xi_\tau \sim (\mathbf{0}, \Omega), \quad \tau = 3, 6, \dots T \tag{2.5}$$

where $\Omega = BB' + ABB'A' + A^2BB'A'^2$. We consider the following moment conditions:

$$C = A^3, \tag{2.6}$$

$$\Omega = BB' + ABB'A' + A^2BB'A'^2 \tag{2.7}$$

[8] Foroni and Marcellino (2014) evaluate analytically the mismatch between the results with the *naive* approach and aggregation techniques. They demonstrate that standard strategy and aggregated technique provides different relations between the structural parameters and the estimated coefficient matrix.

which represent the mapping between the matrices of the high frequency (structural form) and the low frequency model (reduced form).

With B lower triangular, the model is exactly identified: the estimated parameters at quarterly frequency are $n^2 + n(n+1)/2$, where n^2 is the number of distinct elements of C and $n(n+1)/2$ is the number of distinct entries of Ω. In particular, the number of estimated parameters at quarterly frequency are equal to the number of free parameters in the monthly process.
If we need to consider non-recursive structural schemes, at quarterly frequency the number of free parameters is always $n^2 + n(n+1)/2$. However, at monthly frequency the number of parameters which has to be estimated strictly depends on the structure of B, and in particular on the number of free parameters in B. In this discussion we consider the more general situation, specifying B to be full. In this case, the number of high frequency parameters which has to be estimated is $2n^2 > n^2 + n(n+1)/2$, where n^2 are the distinct parameters in A and also in B. Hence, considering the most general case of the non-recursive schemes, the monthly SVAR(1) is not identified.
In recent years, many authors have proposed to use specific characteristics of the data, with the aim of providing further moment conditions to identify and estimate the structural parameters. Lanne and Lutkepohl (2008) consider the idea of matching and testing conventional linear restrictions (especially, of non-recursive identification schemes) and statistical information in order to consider time varying structures in the error covariance matrix[9]. In the situation of possible structural break(s) in the data, we can recover to the solution proposed by Lanne and Lutkepohl (2008).
Assume that we observe a singular structural change in the volatility at time T_b, with $1 < T_b < T$. Then the ("reduced") high frequency model is defined by

$$Y_t^* = AY_{t-1}^* + u_t, \quad u_t \sim \begin{cases} (0, \Sigma_1), & 1 \geq t \geq T_b \\ (0, \Sigma_2), & T_b + 1 \geq t \geq T. \end{cases} \tag{2.8}$$

Given $\Sigma_1 \neq \Sigma_2$, we can rewrite the covariance matrix Σ_1 as BB' and $\Sigma_2 = BVB'$, with $V \neq I$ is a $n \times n$ diagonal matrix. By this way we can easily estimate $n^2 + n(n+1)/2$ parameters for each equation (n^2 free parameters in A, and $n(n+1)/2$ distinct elements - due to symmetry - in the covariance matrix Σ_i, $i = 1, 2$); then the total number of estimated parameters in Eq. (2.8) is $n^2+n(n+1)$. On the other hand, the total number of free parameters in the structural form is $n^2 + n^2 + n$ (n^2 free parameters in A, n^2 structural coefficients in B and n distinct elements in the diagonal covariance matrix V), i.e. equal to the number

[9]Lanne and Lutkepohl (2008) adapt the approach of Klein and Vella (2006) and Rigobon (2003) who shown that identification can be obtained using the additional information of the heteroscedasticity found in the data.

of estimated coefficients in the reduced form. Introducing heteroscedasticity, also the monthly process with non-recursive structural scheme is (exactly) identified. Specifically, we consider to have just detected a singular structural break in the volatility at time T_b, where $1 < T_b < T$. For sake of simplicity, we consider the following assumptions:

ASSUMPTION 1: the high frequency instant in which we observe a structural break in the volatility T_b coincides with the third month of the reference quarter.

ASSUMPTION 2: the change in the error covariance matrix Σ is detectable also with quarterly data.

The second assumption stems from the idea that, even if the heteroscedastic behaviour of the data can be reduced by temporal aggregation, the empirical evidence of change through distinct regimes can not be excluded[10].
After the detection of (two) distinct volatility regimes, at the low frequency we estimate $n^2 + n(n+1)$ parameters, that coincides with the number of the high frequency free parameters ($n^2 + n^2 + n = n^2 + n(n+1)$). In this case, the monthly model specification is given by

$$Y_t^* = AY_{t-1}^* + B\varepsilon_t, \quad \varepsilon_t \sim \begin{cases} (\mathbf{0}, I_3), & \text{if } t = 1, \dots T_b \\ (\mathbf{0}, V), & \text{if } t = T_b + 1, \dots, T \end{cases} \tag{2.9}$$

where $V = diag(v_1, v_2, v_3)$, with quarterly correspondent processes:

$$Y_\tau = CY_{\tau-1} + \xi_\tau, \quad \xi_\tau \sim \begin{cases} (\mathbf{0}, \Omega_1), & \text{if } \tau = 3, 6, \dots T_b \\ (\mathbf{0}, \Omega_2), & \text{if } \tau = T_b + 3, \dots, T \end{cases} \tag{2.10}$$

where $\Omega_1 = BB' + ABB'A' + A^2BB'A'^2$ and $\Omega_2 = BVB' + ABVB'A' + A^2BVB'A'^2$. In this case the mapping between the monthly and the quarterly matrices is given by:

$$C = A^3, \tag{2.11}$$
$$\Omega_1 = BB' + ABB'A' + A^2BB'A'^2, \tag{2.12}$$
$$\Omega_2 = BVB' + ABVB'A' + A^2BVB'A'^2. \tag{2.13}$$

Our idea is to obtain the high frequency parameters using the mapping reported above. Specifically we recover a preliminary estimate of the monthly coefficient

[10]The classical example of empirical applications in macroeconomic framework that has been investigated is represented by the transition from the 'Great Inflation' to the 'Great Moderation' period. See e.g. Boivin and Giannoni (2006) and Bacchiocchi and Fanelli (2015).

matrix A, i.e. $\tilde{A}$, solving the (real) cube root of $\hat{C}$ (indirect estimate); we use $\tilde{A}$ as starting values for the minimum distance estimation of $\hat{A}$, $\hat{B}$ (and $\hat{V}$), given Eq. (2.11)-(2.13).

The real cube of C

The high nonlinearity of the relation $C = A^3$ can be solved using the Jordan canonical form. Higham (2008) summarizes three approaches to evaluate matrix functions: Jordan canonical form, polynomial interpolation and Cauchy integral. From this starting point a wide range of the literature concentrates the attention on pth roots of matrices. In this specific case, the cube root of a square matrix in $\mathbb{R}^{n\times n}$, with no repeated eigenvalues (i.e. non-derogatory matrix), appears as a very specific situation of a more general framework. In the case of VAR coefficient matrix estimation we focus only on solving this specific problem. See APPENDIX A1 for further details on matrix functions, Jordan canonical form and for a more general treatment of this specific problem.
We consider the Jordan decomposition $C = MJM^{-1}$, where M is an $n \times n$ non-singular matrix, and J is the $n \times n$ block diagonal Jordan matrix. The function $C^{1/3} \in \mathbb{R}^{n\times n}$, is defined by

$$\begin{aligned} C^{1/3} &= (MJM^{-1})^{1/3} \\ &= MJ^{1/3}M^{-1} \\ &= Mdiag(J_1(\lambda_1)^{1/3}, \ldots, J_n(\lambda_n)^{1/3})M^{-1}, \end{aligned}$$

where λ_i, $i = 1, \ldots, n$, are the (non-repeated) eigenvalues of C and the diagonal (Jordan) blocks $J_i(\lambda_i)$ of the Jordan matrix J are scalars, $J_i = \lambda_i$, $i = 1, \ldots, n$.

The problem in the estimation of the monthly coefficient matrix A, is that $A = C^{1/3}$ not always has a unique solution, also in the specific case of $C^{1/3} \in \mathbb{R}^{n\times n}$. Starting with the definition of the Jordan decomposition, and considering the example with $n = 3$, we have to evaluate two different cases[11]: (i) the cube root of a matrix C with 3 real distinct eigenvalues, (ii) the cube root of a matrix C with one real eigenvalue and a pair of complex conjugate eigenvalues. In the first case, the solution $C^{1/3} = A \in \mathbb{R}^{3\times 3}$ is unique. In the second case, what we obtain are three distinct solutions, $C^{1/3} = \{A_1, A_2, A_3\} \in \mathbb{R}^{3\times 3}$. In this case, the general idea

[11]In a SVAR with $n > 3$, the number of possible pair of complex conjugate eigenvalues has to be an even-number (or equal to zero). The number h of possible distinct solutions in $\mathbb{R}$ of the cube root of a $n \times n$ matrix is related to the amount k of complex conjugates pairs and, in particular $h = 3^k$. For example, in a VAR(1) with $n = 5$ endogenous variables, where the coefficient matrix C has one real eigenvalue and two pairs of complex conjugates eigenvalues, the number of equivalent solutions of $C^{1/3}$ is $3^2 = 9$.

is to use the estimation of the structural coefficient matrices obtained for each A_i, $i = 1, 2, 3$, in order to identify the "true" set of monthly parameters.

Minimum Distance estimation

Suppose that θ_0 is the $n_\theta \times 1$ vector of (high frequency) parameters of interest, which is known to be function of the $n_\phi \times 1$ (low frequency) parameter vector ϕ_0, with $n_\phi \geq n_\theta$. In particular, for a known continuously differentiable function h,

$$f(\phi_0, \theta_0) = \phi_0 - h(\theta_0) = 0.$$

Let $\hat{\phi}$ be a consistent estimator of ϕ, with asymptotic covariance matrix Ψ.
The minimum distance estimator $\hat{\theta}$ solves the minimization problem:

$$\min_{\theta \in \mathcal{T}} Q(\theta) = \min_{\theta \in \mathcal{T}} \{\hat{\phi} - h(\theta)\}' S \{\hat{\phi} - h(\theta)\}, \tag{2.14}$$

where S is any positive semi-definite symmetric matrix.
In the case of B obtained from a Cholesky decomposition the high frequency parameter vector is given by $\theta = (vec(A)', vech(B)')'$. As described above, we know that the $n^2 + n(n+1)/2 \times 1$ (quarterly) parameter vector $\hat{\phi} = (vec(\hat{C})', vech(\hat{\Omega})')'$ is function of θ. The relationship between the number of monthly and quarterly parameters can be summarized by

$$dim(\hat{\phi}) = n^2 + n(n+1)/2 = n^2 + n(n+1)/2 = dim(\theta).$$

The mapping of Eq. (2.6) and Eq. (2.7) represents the known continuously function h shown in Eq (2.14). As described by Newey and McFadden (1994) the optimal choice of the weighting matrix S is represented by the inverse of the asymptotic covariance matrix estimator $\hat{\Psi}$ of the quarterly estimated vector $\hat{\phi}$[12].
Since $S = \hat{\Psi}^{-1}$, the asymptotic distribution of the classical minimum distance estimator $\hat{\theta}$ is given by $\sqrt{T}(\hat{\theta} - \theta_0) \sim (0, \Theta)$, with

$$\hat{\Theta} = \left(\left(\frac{\partial h(\theta)}{\partial \hat{\theta}'} \right)' \hat{\Psi}^{-1} \left(\frac{\partial h(\theta)}{\partial \hat{\theta}'} \right) \right)^{-1},$$

where

$$\hat{\Psi} = \begin{pmatrix} (\hat{\Xi}^{-1} \otimes \hat{\Omega}) & 0_{n^2 \times n(n+1)/2} \\ 0_{n(n+1)/2 \times n^2} & 2D_n^+(\hat{\Omega} \otimes \hat{\Omega})(D_n^+)' \end{pmatrix},$$

and $D_n^+ = (D_n' D_n)^{-1} D_n'$ is the $(n^2 \times n(n+1)/2)$ Moore-Penrose inverse of the duplication matrix D_n and $\hat{\Xi} = Y_{t-1} Y_{t-1}' / T$.

[12]The definition of the optimal weighting matrix is discussed in detail in APPENDIX A3

With B non-recursive, we consider the quarterly $n^2 + n(n+1) \times 1$ vector of parameters $\hat{\phi} = (vec(\hat{C})', vech(\hat{\Omega}_1)', vech(\hat{\Omega}_2)')'$, and evaluate the relationships between $\hat{\phi}$ and the monthly vector given by $\theta = (vec(A)', vec(B)', diag(V)')'$ of dimension $2n^2 + n \times 1$. What we observe is

$$dim(\hat{\phi}) = n^2 + n(n+1) = 2n^2 + n = dim(\theta)$$

that is, the monthly process is exactly identified. In particular, for the known continuously differentiable function h, defined by the relationships in Eq. (2.11), (2.12) and (2.13), the minimization problem is represented by Eq. (2.14), with optimal weighting matrix S given by the inverse of the estimator $\hat{\Psi}$ of the asymptotic covariance matrix of the estimated $\hat{C}$, $\hat{\Omega}_1$ and $\hat{\Omega}_2$. In particular, the block diagonal matrix $\hat{\Psi}$ is defined by

$$\hat{\Psi} = \begin{pmatrix} \hat{\Psi}_{00} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \hat{\Psi}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \hat{\Psi}_{22} \end{pmatrix},$$

where $\hat{\Psi}_{00} = (\hat{\Xi}^{-1} \otimes \hat{\Omega})$ is the estimated covariance matrix of $vec(\hat{C})$ calculated on the entire period, and $\hat{\Psi}_{11} = 2D_n^+(\hat{\Omega}_1 \otimes \hat{\Omega}_1)(D_n^+)'$ and $\hat{\Psi}_{22} = 2D_n^+(\hat{\Omega}_2 \otimes \hat{\Omega}_2)(D_n^+)'$ are the estimated covariance matrix of the quarterly estimates of the covariance matrices of the residuals, respectively calculated for $\tau = 3, \dots, T_b$ and $\tau = T_b + 3, \dots, T$.

Coming back to the initial values in the computation of the real cube root of $\hat{C}$, when the eigenvalues of the estimated quarterly coefficient matrix C are all (distinct) in the real plane, the solution $\hat{C}^{1/3} = \hat{A} \in \mathbb{R}^{3\times 3}$ is unique, hence also $\hat{B}$ (and $\hat{V}$, in the case of B non-recursive). Otherwise, if $\hat{C}^{1/3} = \{\tilde{A}_1, \tilde{A}_2, \tilde{A}_3\} \in \mathbb{R}^{3\times 3}$, with the minimum distance estimation we obtain $\{\hat{A}_1, \hat{B}_1\}$, $\{\hat{A}_2, \hat{B}_2\}$ and $\{\hat{A}_3, \hat{B}_3\}$ (or $\{\hat{A}_1, \hat{B}_1, \hat{V}_1\}$, $\{\hat{A}_2, \hat{B}_2, \hat{V}_2\}$ and $\{\hat{A}_3, \hat{B}_3, \hat{V}_3\}$), respectively associated to $\{\tilde{A}_1, \tilde{A}_2, \tilde{A}_3\}$.
Then, the procedure follows with the identification of the *"true"* set of matrices among the three sets previously obtained and summarized as follows:

$$\text{B lower triangular} \quad \begin{cases} \hat{\theta}_1 = \{\hat{A}_1, \hat{B}_1\} \\ \hat{\theta}_2 = \{\hat{A}_2, \hat{B}_2\} \\ \hat{\theta}_3 = \{\hat{A}_3, \hat{B}_3\} \end{cases}$$

or

$$\text{B non-recursive} \quad \begin{cases} \hat{\theta}_1 = \{\hat{A}_1, \hat{B}_1, \hat{V}_1\} \\ \hat{\theta}_2 = \{\hat{A}_2, \hat{B}_2, \hat{V}_2\} \\ \hat{\theta}_3 = \{\hat{A}_3, \hat{B}_3, \hat{V}_3\} \end{cases} .$$

The proposed approach considers to chose the set of estimated matrices that provides the minimum value of the minimization function calculated for each of the three cases, i.e. $\{Q(\hat{\theta}_1), Q(\hat{\theta}_2), Q(\hat{\theta}_3)\}$.
Even if the order condition guarantees the identifiability of the monthly parameters, in the estimation of the model some problems could arise, due to the high nonlinearity of the relationships described above in Eq. (2.6) and Eq. (2.7), and in Eq. (2.11), (2.12) and (2.13). The solution considered is to solve the minimization problems, given the relationships:

- *B lower triangular*

$$C = A^3, \tag{2.15a}$$
$$\Omega = \Sigma + A\Sigma A' + A^2\Sigma A'^2, \tag{2.15b}$$

  with $\Sigma = BB'$, and then obtain B with the Cholesky decomposition;

- *B non-recursive*

$$C = A^3, \tag{2.16a}$$
$$\Omega_1 = \Sigma_1 + A\Sigma_1 A' + A^2\Sigma_1 A'^2, \tag{2.16b}$$
$$\Omega_2 = \Sigma_2 + A\Sigma_2 A' + A^2\Sigma_2 A'^2, \tag{2.16c}$$

  with $\Sigma_1 = BB'$, $\Sigma_2 = BVB'$, and B (and V) estimated with a final Minimum Distance estimation step, based on the explicit relationships

$$\Sigma_1 = BB', \tag{2.17a}$$
$$\Sigma_2 = BVB', \tag{2.17b}$$

and weighting matrix S given by the inverse of the asymptotic covariance matrix[13] $\hat{\Theta}^{(1,2)}$ of the estimated $\hat{\theta}^{(1,2)} = (vech(\hat{\Sigma}_1)', vech(\hat{\Sigma}_2)')'$ and obtained at the previous step as

$$\hat{\Theta}^{(1,2)} = \left(\left(\frac{\partial h(\theta^{(1,2)})}{\partial \hat{\theta}^{(1,2)\prime}} \right)' \begin{pmatrix} \hat{\Psi}_{11} & \mathbf{0} \\ \mathbf{0} & \hat{\Psi}_{22} \end{pmatrix}^{-1} \left(\frac{\partial h(\theta^{(1,2)})}{\partial \hat{\theta}^{(1,2)\prime}} \right) \right)^{-1}.$$

The proposed decomposition of the problem simplifies the estimation procedure and removes any possible failure of identification due to high nonlinearity of the relationships.

[13]See APPENDIX A3 for further details about Minimum Distance and Classical Minimum Distance estimation.

2.3.2 Monte Carlo experiments

We consider two different Monte Carlo experiments, with different sample sizes, to evaluate the solutions proposed above either for B lower triangular, or for B nontriangular.
For $M = 1000$ replications, we generate the monthly trivariate SVAR(1) process defined in Eq. (2.2), where Y_t^*, $t = 1\ldots,T$, is the $n \times 1$ vector of the n monthly series, Y_0 is set to $0_{n\times 1}$, A is the $n \times n$ coefficient matrix, B is the $n \times n$ matrix of coefficients of instantaneous shocks. We consider different sample sizes, in particular $T = \{600, 1200, 3600\}$ (months, corresponding to $200, 400$ and 1200 quarters)[14].
For each replication, we apply the filters $\omega(L) = I$ and $\Gamma(L) = (I + AL + A^2L^2)$, to the monthly series Y_t^*. The result is the quarterly variable Y_τ, aggregated *via* point-in-time sampling.
In the case of B lower triangular, the estimation procedure, proposed above, corresponds to:

- obtain the quarterly least squares estimates of $\hat{C}$ and $\hat{\Omega}$, from the quarterly VAR(1) of Eq. (2.5);
- solve the (real) cube root $\hat{C}^{1/3} = \tilde{A}$;
- use $\tilde{A}$ as starting values for the Minimum distance estimation of A and $\Sigma = BB'$, given the mapping of Eq. (2.15);
- use the Cholesky decomposition of $\hat{\Sigma}$ to obtain $\hat{B}$.

Similarly, for the case of B non recursive, we refer to the monthly process in Eq. (2.9), where $V = diag(v_1, v_2, v_3)$, T_b is the break in the process and, for simplicity, we assume that (i) T_b coincides with the third month of the reference quarter, and (ii) $T_b = T/2$. In this case the estimation approach can be summarized by the following steps:

- estimate the quarterly matrices $\hat{C}$, $\hat{\Omega}_i$, with $i = 1, 2$;
- solve the cube root $\tilde{A} = \hat{C}^{1/3}$;
- estimate the monthly vector of parameters $(vec(\hat{A})', vech(\hat{\Sigma}_1)', vech(\hat{\Sigma}_2)')'$, solving the minimization problem in Eq. (2.14), with the mapping described by the system in Eq. (2.16);
- we consider a further Minimum distance step in which we estimate $\hat{B}$ and $\hat{V}$ from the relationships in Eq. (2.17).

[14]Even if the seconds sample sizes seem to be unreasonable, we shown these results to highlight how the accuracy of the estimates improves working with large samples.

We report below two representative examples. Referring to B recursive, we consider the monthly DGP

$$Y_t^* = \underbrace{\begin{pmatrix} 0.500 & 0.320 & 0.100 \\ 0.200 & 0.250 & 0.050 \\ 0.100 & 0.400 & 0.800 \end{pmatrix}}_{A} Y_{t-1}^* + \underbrace{\begin{pmatrix} 0.500 & 0 & 0 \\ 0.300 & 0.400 & 0 \\ 0.100 & 0.250 & 0.600 \end{pmatrix}}_{B} \varepsilon_t, \qquad \varepsilon_t \sim (0, I) \tag{2.18}$$

The monthly SVAR(1) process in Eq. (2.18) generates the quarterly point-in-time VAR(1) process defined by:

$$Y_\tau = \underbrace{\begin{pmatrix} 0.233 & 0.232 & 0.163 \\ 0.114 & 0.115 & 0.081 \\ 0.263 & 0.448 & 0.579 \end{pmatrix}}_{C} Y_{\tau-1} + \xi_\tau, \qquad \xi_\tau \sim \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \underbrace{\begin{pmatrix} 0.501 & 0.282 & 0.390 \\ 0.282 & 0.320 & 0.314 \\ 0.390 & 0.314 & 1.23 \end{pmatrix}}_{\Omega} \right). \tag{2.19}$$

The representative design of Eq. (2.18) and the DGP matrices of the case of B non-recursive, are reported in the first columns of Table 2.1 and Table 2.2 respectively, with the related Monte Carlo results. We report also the impulse response function for both the Monte Carlo experiments, with $T = 3600$, in Figure 2.2 and Figure 2.3 - 2.4. We depict the response of the i-th variable to a shock on the j-th variable as $(j \to i)$ i.e. (*impulse* $\to$ *response*). The biases of the estimates are measured referring to two quantities: first, the sum of the differences, in absolute value, of the DGP values and the Monte Carlo estimates (in the Tables as $bias(1)$), and $bias(2)$, calculated as $bias(1)$/(number of estimated elements).

In Table 2.1 and Table 2.2 we report the true value of the parameters in the first column, and for each $T = \{600, 1200, 3600\}$ we show the estimates of the high frequency parameters, the Monte Carlo and the empirical standard errors (as "MCs.e." and "s.e." respectively), these last obtained from the square root of the diagonal entries of the estimated covariance matrix of the minimum distance estimates[15]. An enhancement in the performances of the proposed approach have been observed increasing the sample size. Also the standard errors of the estimates get better and gradually decreases, passing from $T = 600$ high frequency periods to $T = 3600$.
However, the results assess the difficulty of the Minimum Distance estimation procedure to treat the high nonlinearity of the examined relationships. In the next

[15] A possible alternative, which we don't consider in the Monte Carlo experiments, is represented by the bootstrap procedure.

section we reduce the sample size and we compare the Minimum Distance estimation and the state space methodology, in order to evaluate both the mixed frequency estimation techniques in term of structural analysis.

	T_m=600			T_m=1200			T_m=3600		
$vec(A)$	$vec(\hat{A})$	MCs.e.	s.e.	$vec(\hat{A})$	MCs.e	s.e.	$vec(\hat{A})$	MCs.e.	s.e.
0.500	0.466	0.04	0.03	0.486	0.04	0.02	0.490	0.03	0.02
0.200	0.304	0.09	0.06	0.258	0.06	0.04	0.233	0.03	0.03
0.100	0.082	0.07	0.05	0.075	0.05	0.03	0.082	0.04	0.02
0.320	0.332	0.06	0.03	0.322	0.06	0.02	0.328	0.03	0.03
0.250	0.120	0.09	0.05	0.172	0.07	0.04	0.204	0.04	0.04
0.400	0.450	0.06	0.05	0.441	0.03	0.03	0.428	0.03	0.02
0.100	0.110	0.04	0.01	0.104	0.03	0.01	0.102	0.02	0.00
0.050	0.037	0.07	0.02	0.045	0.06	0.01	0.047	0.04	0.01
0.800	0.794	0.06	0.01	0.799	0.03	0.01	0.800	0.03	0.00
bias(1)	0.377	-	-	0.229	-	-	0.149	-	-
bias(2)	0.042	-	-	0.025	-	-	0.017	-	-
$vech(B)$	$vech(B)$	MCs.e.	s.e.	$vech(B)$	MCs.e.	s.e.	$vech(B)$	MCs.e.	s.e
0.500	0.492	0.06	0.01	0.494	0.04	0.01	0.496	0.03	0.01
0.300	0.307	0.06	0.02	0.314	0.06	0.01	0.299	0.05	0.01
0.100	0.123	0.09	0.03	0.106	0.07	0.02	0.088	0.06	0.02
0.400	0.361	0.04	0.01	0.368	0.03	0.01	0.386	0.02	0.01
0.250	0.278	0.14	0.02	0.297	0.14	0.02	0.272	0.13	0.02
0.600	0.554	0.13	0.02	0.574	0.14	0.01	0.589	0.13	0.01
bias(1)	0.152	-	-	0.129	-	-	0.063	-	-
bias(2)	0.025	-	-	0.022	-	-	0.010	-	-

Table 2.1: Monte Carlo results obtained for point-in-time sampling of a SVAR(1) with B 'Cholesky based'. In the first column we report the elements of the population matrices A and B defined in Eq. (2.18). We consider $T = 300, 600, 1200$, and we evaluate the results with two measures of bias: $bias(1)$ is calculated as the sum of the differences in absolute value of the "true" value of the elements and the Monte Carlo estimates; $bias(2)$ is calculated as $bias(1)$/(number of elements).

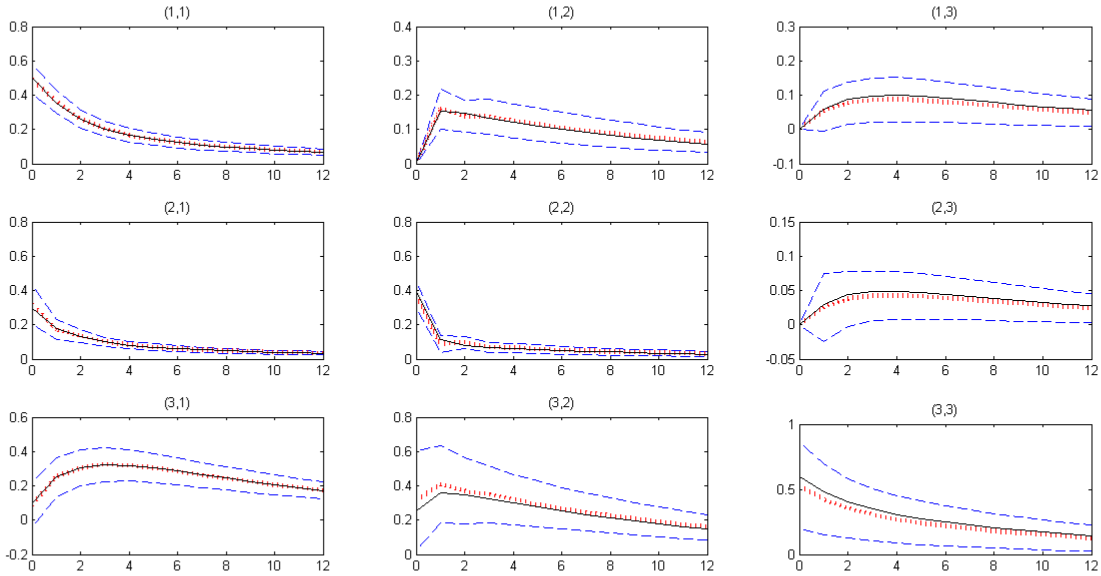


Figure 2.2: Impulse responses from Monte Carlo experiment reported in Table 2.1, with sample size $T = 3600$. The red-dotted line corresponds to the population responses, the blue-dashed lines refer to the confidence bounds (calculated as two-standar error bounds) of the IRFs calculated with the proposed Minimum distance approach (black line)

.

	T_m=600			T_m=1200			T_m=3600		
$vec(A)$	$vec(\hat{A})$	MCs.e.	s.e.	$vec(\hat{A})$	MCs.e	s.e.	$vec(\hat{A})$	MCs.e.	s.e.
0.500	0.520	0.05	0.08	0.517	0.05	0.07	0.512	0.04	0.06
0.100	0.081	0.06	0.10	0.076	0.05	0.08	0.078	0.04	0.07
0.050	0.062	0.05	0.08	0.068	0.04	0.07	0.068	0.04	0.06
0.200	0.129	0.09	0.18	0.141	0.08	0.16	0.153	0.08	0.15
0.350	0.434	0.13	0.21	0.428	0.11	0.20	0.421	0.11	0.19
0.400	0.347	0.11	0.16	0.347	0.09	0.15	0.347	0.09	0.15
-0.100	-0.069	0.05	0.09	-0.073	0.04	0.08	-0.079	0.04	0.07
0.250	0.210	0.06	0.10	0.215	0.05	0.10	0.218	0.05	0.08
0.700	0.727	0.05	0.08	0.724	0.05	0.08	0.723	0.04	0.07
bias(1)	0.367	-	-	0.279	-	-	0.220	-	-
bias(2)	0.041	-	-	0.031	-	-	0.024	-	-
$vec(B)$	$vec(\hat{B})$	MCs.e.	s.e.	$vec(\hat{B})$	MCs.e.	s.e.	$vec(\hat{B})$	MCs.e.	s.e
0.250	0.233	0.17	0.35	0.237	0.16	0.27	0.245	0.13	0.14
0.150	0.154	0.10	0.29	0.149	0.10	0.24	0.127	0.08	0.14
0.100	0.170	0.11	0.36	0.163	0.10	0.30	0.160	0.08	0.17
0.500	0.450	0.14	0.27	0.459	0.14	0.21	0.491	0.11	0.09
0.300	0.249	0.08	0.18	0.259	0.08	0.14	0.282	0.06	0.06
-0.230	-0.199	0.07	0.15	-0.209	0.05	0.08	-0.210	0.04	0.03
0.200	0.179	0.10	0.17	0.191	0.08	0.10	0.196	0.05	0.04
-0.250	-0.241	0.07	0.11	-0.244	0.05	0.05	-0.249	0.03	0.02
0.430	0.364	0.10	0.24	0.381	0.07	0.14	0.392	0.04	0.08
bias(1)	0.304	-	-	0.235	-	-	0.208	-	-
bias(2)	0.034	-	-	0.026	-	-	0.023	-	-
$diag(V)$	$diag(\hat{V})$	MCs.e.	s.e.	$diag(\hat{V})$	MCs.e.	s.e.	$diag(\hat{V})$	MCs.e.	s.e
1.600	1.438	0.68	1.27	1.463	0.58	0.80	1.472	0.45	0.52
2.400	2.391	0.62	0.95	2.380	0.50	0.59	2.384	0.34	0.31
0.500	0.487	0.17	0.45	0.491	0.14	0.27	0.513	0.13	0.15
bias(1)	0.185	-	-	0.166	-	-	0.157	-	-
bias(2)	0.062	-	-	0.055	-	-	0.052	-	-

Table 2.2: Monte Carlo results obtained for point-in-time sampling of a SVAR(1) with B non-recursive. In the first column we report the elements of the population matrices A, B and V. We consider $T = 300, 600, 1200$, and we evaluate the results with two measures of bias: $bias(1)$ is calculated as the sum of the differences in absolute value of the "true" value of the elements and the Monte Carlo estimates; $bias(2)$ is calculated as $bias(1)/(\text{number of elements})$.

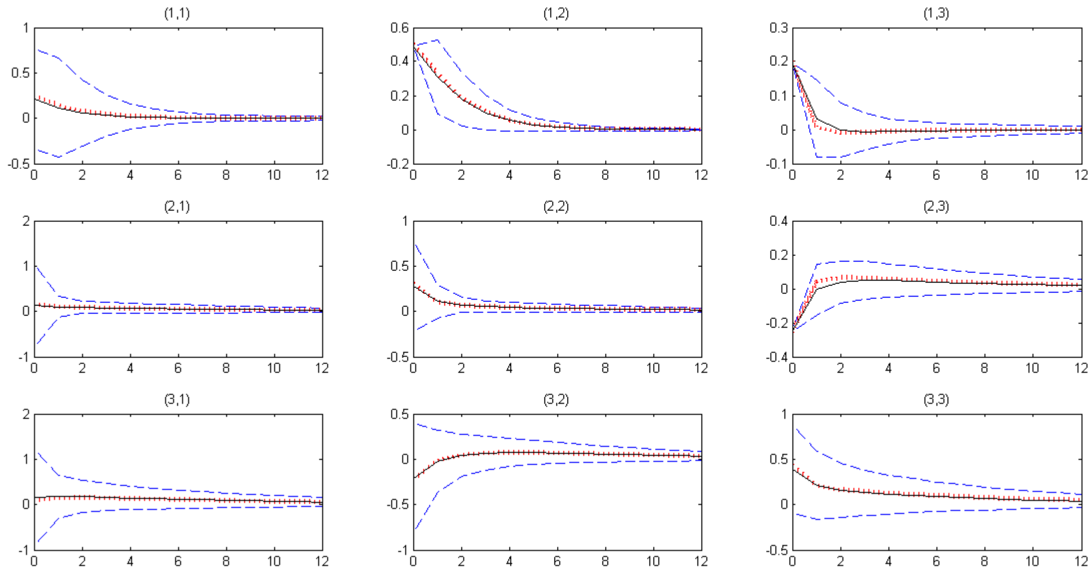


Figure 2.3: Impulse responses from Monte Carlo experiment reported in Table 2.2, with sample size $T = 3600$ and $t = 1, \ldots, T_b$. The red-dotted line corresponds to the population responses, the blue-dashed lines refer to the confidence bounds (calculated as two-standar error bounds) of the IRFs calculated with the proposed Minimum distance approach (black line)

.

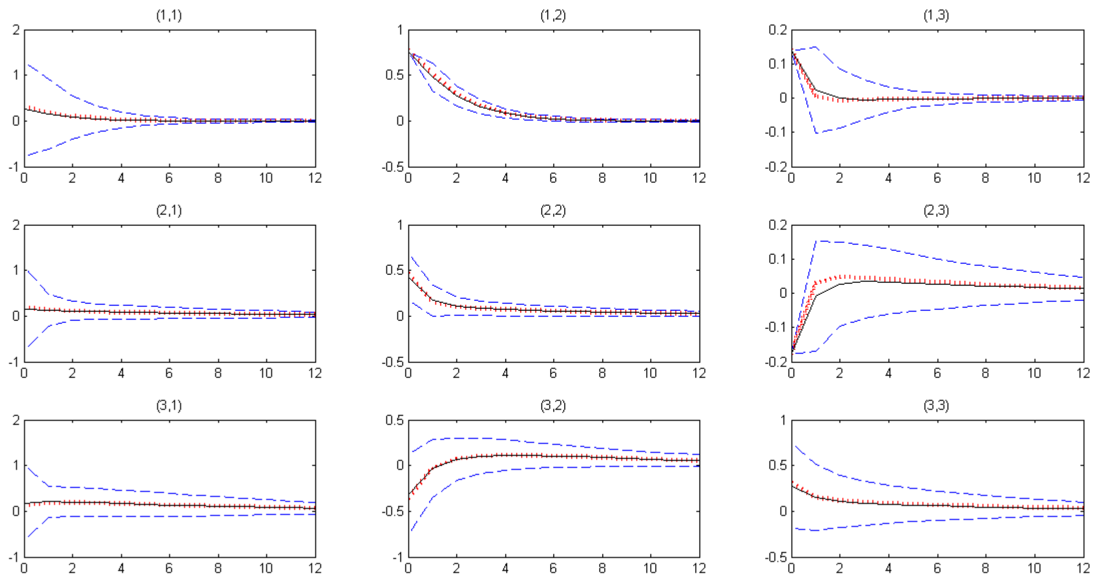


Figure 2.4: Impulse responses from Monte Carlo experiment reported in Table 2.2, with sample size $T = 3600$ and $t = T_b + 1, \ldots, T$. The red-dotted line corresponds to the population responses, the blue-dashed lines refer to the confidence bounds (calculated as two-standar error bounds) of the IRFs calculated with the proposed Minimum distance approach (black line)

.

2.3.3 Comparison with state space estimation method

In this section we consider four different Monte Carlo experiments to compare the approach proposed in this discussion with results obtained with the state space approach, both in term of parameters estimation and impulse responses. In order to reduce the computational burden, in the following Monte Carlo exercises we refer only to bivariate DGPs.
Specifically, the general design corresponds to a high frequency bivariate SVAR(1), defined by

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \underbrace{\begin{pmatrix} \rho_1 & \delta_{1,2} \\ \delta_{2,1} & \rho_2 \end{pmatrix}}_{A} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \underbrace{\begin{pmatrix} b_1 & b_{1,2} \\ b_{2,1} & b_2 \end{pmatrix}}_{B} \begin{pmatrix} u_{y,t} \\ u_{x,t} \end{pmatrix}, \qquad t = 1, \ldots, T, \tag{2.20}$$

where y_t is the variable observed only in the third month of the reference quarter, x_t is the high frequency variable and

$$\underbrace{\begin{pmatrix} b_1 & b_{1,2} \\ b_{2,1} & b_2 \end{pmatrix}}_{B} \begin{pmatrix} u_{y,t} \\ u_{x,t} \end{pmatrix} = \begin{pmatrix} e_{y,t} \\ e_{x,t} \end{pmatrix} \sim \mathcal{N}_2(0, \Sigma). \tag{2.21}$$

The vector $(u'_{y,t}, u'_{x,t})'$ is sampled from a standard normal distribution, and $t = 1, \ldots, T$ is the (monthly) time index, and $T = 300$ months (corresponding to 100 quarters, 25 years). In the firsts two experiments, $B = I_2$, leading $\Sigma = I_2$; in the other Monte Carlo designs we refer to B lower triangular matrix, obtained from the Cholesky decomposition of Σ (i.e. $b_{1,2} = 0$). In particular, the population matrix A is taken from

$$\left\{ \begin{pmatrix} 0.75 & 0.1 \\ 0.2 & 0.7 \end{pmatrix}, \begin{pmatrix} 0.85 & 0.1 \\ 0.2 & 0.8 \end{pmatrix} \right\}$$

and B from

$$\left\{ \begin{pmatrix} 1.0 & 0.0 \\ 0.0 & 1.0 \end{pmatrix}, \begin{pmatrix} 0.6 & 0.0 \\ 0.3 & 0.4 \end{pmatrix} \right\}.$$

Since the proposed approach works on the eigenvalues of A, we don't pay attention to the similarity of the entries of two matrices in first sample, but we consider their eigenvalues: in the first case the eigenvalues are $(0.8686, 0.5814)$, while the second choice of A allows us to evaluate those cases in which the larger eigenvalue is close to the unity, i.e. $(0.9686, 0.6814)$.
In the first part of the experiments, we evaluate the performances with both the mixed frequency estimation techniques, after 1000 replications, with the measure *bias(1)*, considered above. In the second part of the experiment, we pay attention to the IRFs. We estimate:

- the high frequency VAR(1) parameters with ordinary least squares (benchmark), since the simulation allows us to know the latent high frequency process underlying the quarterly variable y_t;
- the mixed frequency model, using the state space approach;
- the mixed frequency model, estimating the high frequency parameters with the Minimum Distance technique.

To make shocks orthogonal, we consider the Cholesky decomposition of the residual covariance matrices, in each approach. The forecast horizons are set to one year after the period of the impact of the shock. Following Foroni and Marcellino (2016), for $M = 1000$ replication, we computed the squared error between the impulse responses obtained for each estimation approach (monthly, state space and minimum distance), and the true impulse responses (available, since the knowledge of the population matrices). For each horizon, we calculate the means across replications of the squared errors (MSE) and we calculate the ratios

$$\frac{RMSE_{MF,SS}}{RMSE_{Mon}} \quad \text{and} \quad \frac{RMSE_{MF,MD}}{RMSE_{Mon}},$$

where $RMSE_{MF,SS}$ is the root mean square error (RMSE) referred to the IRFs obtained with the state space approach, $RMSE_{MF,MD}$ is obtained with the minimum distance estimation, and $RMSE_{Mon}$ is the RMSE calculated with the monthly estimates. By this way, we are able to compare the mixed frequency procedures with the monthly (benchmark) estimation. The RMSE ratios can be seen as a simple (non-exhaustive) indication of the loss of information that we obtain using mixed frequency instead of the high frequency data[16].
What we expect to see are ratios greater then one (i.e. the monthly frequency allows estimation biases smaller then those obtained with the mixed frequency techniques). The higher then one the ratio, the less unreliable the mixed frequency responses are.
Following the considerations of Foroni and Marcellino (2016), in this discussion we are also interested in showing how the mixed frequency estimation procedures alleviate the temporal aggregation bias. Hence, we compare the aggregated mixed frequency responses and low frequency IRFs.

Results

In the first part of the comparison we focus on the estimated parameters. Both the two approaches provide Monte Carlo estimates with associated small biases

[16]See Koelbl, Braumann, Felsenstein and Deistler (2016).

and moderated standard errors. Therefore, we don't observe a unique response on the goodness of one of the estimation methodology over the other.

In the second kind of comparison, referring to the IRFs, the results support the considerations available in the literature: the use of the monthly information in the estimation of the parameters and in the analyses of the impulse responses, guarantees a reduction of the estimation bias. Hence, since the minimum distance is a parameter-driven estimation approach, and it doesn't involve the high frequency information, the state space procedure yields more reliable results than those obtained with the proposed approach. All the ratios calculated for the impulse responses obtained from the state space approach are (as expected) greater then one, but in most of the cases, smaller then the ratios of Minimum Distance estimation.
On the other hand, if Foroni and Marcellino (2016) highlight the hardness of capturing (correctly) the responses of the low frequency variable to shock to the same variable[17], the use of the minimum distance estimation approach can lead some interesting results: in both the experiments with $B \neq I$ (the second half part of Table 2.4 and Table 2.5), we obtain some enhancements of the estimates. In particular, "minimum distance-RMSE-ratios" smaller then the state space counterparts, can be observed for the responses of the low frequency and the high frequency variables, to a shock on the low frequency (in Tables 2.4 and 2.5, *"y to y"* and *"y to x"*).
A further aspect of the problem can be evaluated if we are interested in assessing the bias due to temporal aggregation. As pointed out before, when we assume that the DGP is at monthly frequency, the *naive* estimation approach lead to incorrect responses, and unreliable interpretations. In this case the aggregation bias is easily denoted the last part of both Table 2.6 and Table 2.7: the RMSE ratios calculated comparing the quarterly and the aggregated monthly responses are quite higher then the ratios calculated with both the mixed frequency approaches.

[17]Even if Foroni and Marcellino (2016) consider the comparison of IRFs only after the aggregation, for the point-in-time these considerations are evaluable equivalently.

Design: vec(A)'=(0.75 , 0.20 , 0.10, 0.70)					
vech(B)'=(1.00 , 0.00 , 1.00)			vech(B)'=(0.60 , 0.30 , 0.40)		
TRUE	St.Sp.	Min.Dist.	TRUE	St.Sp	Min.Dist.
$vec(A)$	$vec(\hat{A})$	$vec(\hat{A})$	$vec(A)$	$vec(\hat{A})$	$vec(\hat{A})$
0.75	0.74	0.74	0.75	0.70	0.73
	(0.06)	(0.06)		(0.12)	(0.08)
0.20	0.21	0.22	0.20	0.21	0.23
	(0.05)	(0.07)		(0.07)	(0.07)
0.10	0.1	0.11	0.10	0.14	0.12
	(0.06)	(0.05)		(0.11)	(0.07)
0.70	0.69	0.67	0.70	0.69	0.67
	(0.04)	(0.06)		(0.07)	(0.07)
bias(1)	0.03	0.08	bias(1)	0.12	0.09
$vech(BB')$	$vech(\hat{B}\hat{B}')$	$vech(\hat{B}\hat{B}')$	$vech(BB)$	$vech(\hat{B}\hat{B}')$	$vech(\hat{B}\hat{B}')$
1.00	1.01	1.02	0.36	0.37	0.37
	(0.14)	(0.12)		(0.06)	(0.04)
0.00	0.00	-0.01	0.18	0.18	0.18
	(0.11)	(0.08)		(0.03)	(0.02)
1.00	0.99	0.99	0.25	0.25	0.25
	(0.09)	(0.12)		(0.02)	(0.03)
bias(1)	0.02	0.04	bias(1)	0.01	0.01

Design: vec(A)'=(0.85 , 0.20 , 0.10, 0.80)					
vech(B)'=(1.00 , 0.00 , 1.00)			vech(B)'=(0.60 , 0.30 , 0.40)		
$vec(A)$	$vec(\hat{A})$	$vec(\hat{A})$	$vec(A)$	$vec(\hat{A})$	$vec(\hat{A})$
0.85	0.84	0.84	0.85	0.82	0.84
	(0.05)	(0.03)		(0.08)	(0.05)
0.2	0.21	0.21	0.20	0.21	0.21
	(0.04)	(0.03)		(0.06)	(0.04)
0.1	0.1	0.11	0.10	0.12	0.10
	(0.04)	(0.03)		(0.07)	(0.04)
0.8	0.79	0.79	0.80	0.79	0.79
	(0.03)	(0.03)		(0.05)	(0.04)
bias(1)	0.03	0.05	bias(1)	0.06	0.04
$vech(BB')$	$vech(\hat{B}\hat{B}')$	$vech(\hat{B}\hat{B}')$	$vech(BB')$	$vech(\hat{B}\hat{B}')$	$vech(\hat{B}\hat{B}')$
1.00	1.01	1.02	0.36	0.37	0.36
	(0.17)	(0.10)		(0.06)	(0.03)
0.00	0.00	-0.01	0.18	0.18	0.18
	(0.11)	(0.07)		(0.03)	(0.03)
1.00	0.99	1.00	0.25	0.25	0.25
	(0.08)	(0.09)		(0.02)	(0.02)
bias(1)	0.02	0.03	bias(1)	0.01	0.01

Table 2.3: Estimates from two bivariate SVAR(1) data generating processes, with $T = 300$ high frequency instants. We report the DGP values in the first column, the estimates obtained with state space approach, and the minimum distance estimates (both with related empirical standard errors). In both the experiments of the first half part of the table $vec(A) = (0.75, 0.20, 0.10, 0.70)'$: on the left $B = I$ and on the right part $vech(B) = (0.60, 0.30, 0.40)'$. In the second half part of the table $vec(A) = (0.85, 0.20, 0.10, 0.80)'$, with $B = I$ on the left and $vech(B) = (0.60, 0.30, 0.40)'$ on the right.

Design:	$vec(A) = (0.75, 0.20, 0.10, 0.70)'$, $vech(B) = (1.0, 0.0, 1.0)$							
	St.Sp vs monthly				Min.Dis vs monthly			
h (months)	y to y	x to y	y to x	x to x	y to y	x to y	y to x	x to x
0	2.10	1.98	NaN	1.11	2.22	2.18	NaN	2.34
1	1.26	1.44	1.40	1.13	1.85	2.20	2.12	2.28
2	1.25	1.24	1.34	1.09	1.35	1.48	1.53	1.57
3	1.24	1.16	1.30	1.09	1.29	1.36	1.44	1.44
4	1.20	1.12	1.27	1.10	1.23	1.27	1.37	1.35
5	1.17	1.10	1.25	1.12	1.19	1.22	1.34	1.31
6	1.15	1.09	1.23	1.13	1.16	1.19	1.31	1.28
7	1.13	1.08	1.21	1.14	1.14	1.17	1.29	1.27
8	1.12	1.07	1.20	1.15	1.13	1.16	1.27	1.25
9	1.11	1.06	1.19	1.15	1.12	1.14	1.25	1.24
10	1.10	1.06	1.18	1.15	1.11	1.13	1.24	1.23
11	1.09	1.06	1.17	1.15	1.10	1.13	1.23	1.22
12	1.09	1.06	1.16	1.15	1.10	1.12	1.22	1.22

Design:	$vec(A) = (0.75, 0.20, 0.10, 0.70)'$, $vech(B) = (0.6, 0.3, 0.4)$							
	St.Sp vs monthly				Min.Dis vs monthly			
h (months)	y to y	x to y	y to x	x to x	y to y	x to y	y to x	x to x
0	2.29	1.62	NaN	1.58	2.05	2.01	NaN	2.45
1	2.18	1.32	2.71	1.34	1.41	1.82	1.67	2.17
2	1.39	1.12	1.63	1.18	1.13	1.23	1.29	1.45
3	1.24	1.08	1.43	1.16	1.09	1.12	1.23	1.32
4	1.15	1.07	1.33	1.16	1.07	1.08	1.20	1.23
5	1.11	1.07	1.27	1.16	1.06	1.06	1.18	1.19
6	1.08	1.06	1.24	1.16	1.06	1.06	1.17	1.17
7	1.07	1.05	1.22	1.16	1.06	1.06	1.17	1.16
8	1.06	1.04	1.20	1.15	1.06	1.07	1.16	1.15
9	1.05	1.04	1.19	1.15	1.06	1.07	1.16	1.15
10	1.05	1.04	1.18	1.15	1.06	1.07	1.15	1.14
11	1.04	1.03	1.17	1.14	1.07	1.07	1.15	1.14
12	1.04	1.03	1.16	1.14	1.07	1.08	1.15	1.14

Table 2.4: In the Table we show the RMSE ratios calculated for the evaluation of the Impulse Response Functions for the two mixed frequency procedures. We compare (i) the RMSE for the high frequency IRFs obtained from state space and minimum distance estimation, with (ii) the (benchmark) RMSEs obtained from the estimates the high frequency parameters, once we assume that all the variables can be observed: values near to one, indicate that the mixed frequency impulse responses are close to the monthly impulse responses. In both the fourth and the eighth column, the value "NaN" is due to the Cholesky decomposition (i.e. $b_{1,2} = 0$): the denominator of the ratios is (correctly) equal to 0 at $h = 0$.

Design:	$vec(A) = (0.85, 0.20, 0.10, 0.80)'$, $vech(B) = (1.0, 0.0, 1.0)$							
	St.Sp vs monthly				Min.Dis vs monthly			
h	y to y	x to y	y to x	x to x	y to y	x to y	y to x	x to x
0	1.98	1.89	NaN	1.09	2.03	2.04	NaN	1.92
1	1.23	1.48	1.33	1.16	1.39	1.53	1.66	1.52
2	1.17	1.27	1.30	1.14	1.27	1.37	1.45	1.36
3	1.17	1.19	1.27	1.13	1.25	1.30	1.38	1.31
4	1.17	1.15	1.25	1.13	1.22	1.25	1.34	1.27
5	1.16	1.13	1.23	1.13	1.20	1.22	1.31	1.26
6	1.15	1.11	1.22	1.14	1.18	1.19	1.29	1.24
7	1.13	1.10	1.20	1.15	1.16	1.17	1.27	1.23
8	1.12	1.09	1.19	1.15	1.15	1.16	1.25	1.22
9	1.11	1.09	1.18	1.15	1.13	1.14	1.24	1.21
10	1.10	1.08	1.17	1.15	1.12	1.13	1.23	1.20
11	1.09	1.07	1.16	1.14	1.11	1.12	1.21	1.20
12	1.08	1.07	1.15	1.14	1.10	1.11	1.20	1.19

Design:	$vec(A) = (0.85, 0.20, 0.10, 0.80)'$, $vech(B) = (0.6, 0.3, 0.4)$							
	St.Sp vs monthly				Min.Dis vs monthly			
h	y to y	x to y	y to x	x to x	y to y	x to y	y to x	x to x
0	2.18	1.57	NaN	1.55	1.89	1.89	NaN	1.95
1	1.71	1.26	2.26	1.30	1.29	1.46	1.35	1.44
2	1.29	1.09	1.52	1.19	1.16	1.27	1.30	1.29
3	1.19	1.06	1.37	1.17	1.12	1.17	1.26	1.24
4	1.14	1.06	1.30	1.17	1.10	1.12	1.24	1.22
5	1.11	1.05	1.26	1.17	1.08	1.09	1.23	1.21
6	1.08	1.05	1.24	1.17	1.07	1.07	1.22	1.20
7	1.07	1.05	1.22	1.17	1.06	1.06	1.21	1.19
8	1.06	1.05	1.21	1.17	1.05	1.05	1.21	1.19
9	1.05	1.05	1.20	1.17	1.04	1.05	1.20	1.19
10	1.04	1.04	1.19	1.17	1.04	1.04	1.20	1.19
11	1.04	1.04	1.19	1.17	1.04	1.04	1.19	1.18
12	1.03	1.04	1.18	1.17	1.03	1.04	1.19	1.18

Table 2.5: In the Table we show the RMSE ratios calculated for the evaluation of the Impulse Response Functions for the two mixed frequency procedures. We compare (i) the RMSE for the high frequency IRFs obtained from state space and minimum distance estimation, with (ii) the (benchmark) RMSEs obtained from the estimates the high frequency parameters, once we assume that all the variables can be observed: values near to one, indicate that the mixed frequency impulse responses are close to the monthly impulse responses. In both the fourth and the eighth column, the value “NaN” is due to the Cholesky restriction.

	$vec(A) = (0.75, 0.20, 0.10, 0.70)'$, $vech(B) = (1.00, 0.00, 1.00)'$				$vec(A) = (0.75, 0.20, 0.10, 0.70)'$, $vech(B = (0.60, 0.30, 0.40)'$			
	State Space vs monthly							
h (quarters)	y to y	x to y	y to x	x to x	y to y	x to y	y to x	x to x
0	2.10	1.98	NaN	1.11	2.29	1.62	NaN	1.58
1	1.24	1.16	1.30	1.09	1.24	1.08	1.43	1.16
2	1.15	1.09	1.23	1.13	1.08	1.06	1.24	1.16
3	1.11	1.06	1.19	1.15	1.05	1.04	1.19	1.15
4	1.09	1.06	1.16	1.15	1.04	1.03	1.16	1.14
5	1.08	1.05	1.15	1.40	1.03	1.02	1.16	1.14
6	1.08	1.04	1.14	1.15	1.03	1.02	1.15	1.12
7	1.06	1.04	1.13	1.13	1.03	1.01	1.15	1.11
8	1.05	1.02	1.14	1.13	1.02	1.01	1.14	1.11
	Minimum Distance vs monthly							
0	2.22	2.18	NaN	2.34	2.05	2.01	NaN	2.45
1	1.29	1.36	1.44	1.44	1.09	1.12	1.23	1.32
2	1.16	1.19	1.31	1.28	1.06	1.06	1.17	1.17
3	1.12	1.14	1.25	1.24	1.06	1.07	1.16	1.15
4	1.10	1.12	1.22	1.22	1.07	1.08	1.15	1.14
5	1.08	1.11	1.23	1.21	1.07	1.07	1.15	1.13
6	1.09	1.11	1.22	1.19	1.07	1.05	1.13	1.12
7	1.80	1.09	1.20	1.17	1.06	1.04	1.12	1.13
8	1.60	1.07	1.18	1.15	1.06	1.02	1.11	1.10
	quarterly vs monthly							
0	9.74	6.22	NaN	7.84	11.04	10.59	NaN	6.98
1	3.81	4.27	2.14	2.81	4.30	5.29	1.61	2.10
2	2.95	3.11	1.92	2.20	3.03	3.36	1.54	1.69
3	2.56	2.64	1.84	1.98	2.54	2.70	1.54	1.60
4	2.36	2.41	1.80	1.87	2.32	2.41	1.55	1.58
5	2.25	2.29	1.77	1.83	2.21	2.26	1.56	1.58
6	2.19	2.22	1.76	1.80	2.15	2.19	1.57	1.59
7	2.16	2.18	1.75	1.79	2.11	2.14	1.58	1.59
8	2.14	2.16	1.75	1.79	2.09	2.11	1.59	1.61

Table 2.6: In the Table we report the RMSE ratios for the evaluation of the aggregated IRFs, with poin-in-time-aggregation scheme. We refer to the Monte Carlo designs: (1) $vec(A) = (0.75, 0.20, 0.10, 0.70)'$, $vech(B) = (1.00, 0.00, 1.00)'$ (on the left), and (2) $vec(A) = (0.75, 0.20, 0.10, 0.70)'$, $vech(B) = (0.60, 0.30, 0.40)'$ (on the right). For each (quarterly) horizon, we compare (i) the aggregated Impulse Response Functions of both the mixed frequency procedures, and (ii) the quarterly IRFs (*naive* approach), relative to the monthly IRFs. Values near to one, indicate that the mixed frequency (or the quarterly frequency) impulse responses are close to the monthly impulse responses. In both the fourth and the eighth column, the value "NaN" is due to the Cholesky restriction.

	$vec(A) = (0.85, 0.20, 0.10, 0.80)'$, $vech(B) = (1.00, 0.00, 1.00)'$				$vec(A) = (0.85, 0.20, 0.10, 0.80)'$, $vech(B = (0.60, 0.30, 0.40)'$			
	State Space vs monthly							
h (quarters)	y to y	x to y	y to x	x to x	y to y	x to y	y to x	x to x
0	1.98	1.89	NaN	1.09	2.18	1.57	NaN	1.55
1	1.17	1.19	1.27	1.13	1.19	1.06	1.37	1.17
2	1.15	1.11	1.22	1.14	1.08	1.05	1.24	1.17
3	1.11	1.09	1.18	1.15	1.05	1.05	1.20	1.17
4	1.08	1.07	1.15	1.14	1.03	1.04	1.18	1.17
5	1.07	1.07	1.15	1.15	1.04	1.04	1.17	1.17
6	1.06	1.06	1.15	1.14	1.03	1.03	1.19	1.17
7	1.05	1.05	1.11	1.13	1.02	1.02	1.18	1.16
8	1.04	1.04	1.11	1.12	1.02	1.02	1.16	1.16
	Minimum Distance vs monthly							
0	2.03	2.04	NaN	1.92	1.89	1.89	NaN	1.95
1	1.25	1.30	1.38	1.31	1.12	1.17	1.26	1.24
2	1.18	1.19	1.29	1.24	1.07	1.07	1.22	1.20
3	1.13	1.14	1.24	1.21	1.04	1.05	1.20	1.19
4	1.10	1.11	1.20	1.19	1.03	1.04	1.19	1.18
5	1.06	1.06	1.19	1.18	1.03	1.03	1.19	1.18
6	1.05	1.05	1.16	1.17	1.03	1.03	1.19	1.17
7	1.04	1.04	1.15	1.17	1.02	1.03	1.18	1.18
8	1.04	1.03	1.12	1.15	1.02	1.02	1.16	1.17
	quarterly vs monthly							
0	12.88	7.25	NaN	10.43	14.99	13.25	NaN	9.45
1	5.83	6.29	2.40	4.03	6.88	8.31	1.93	2.85
2	4.50	4.95	2.20	2.80	4.85	5.56	1.82	2.03
3	3.83	4.07	2.07	2.33	3.82	4.23	1.76	1.84
4	3.35	3.51	1.96	2.09	3.19	3.45	1.72	1.76
5	3.00	3.13	1.88	1.95	2.78	2.96	1.68	1.71
6	2.76	2.86	1.82	1.86	2.52	2.64	1.66	1.67
7	2.58	2.66	1.77	1.79	2.34	2.42	1.63	1.65
8	2.46	2.51	1.74	1.75	2.22	2.27	1.61	1.63

Table 2.7: In the Table we report the RMSE ratios for the evaluation of the aggregated IRFs, with poin-in-time-aggregation scheme. We refer to the Monte Carlo designs: (1) $vec(A) = (0.85, 0.20, 0.10, 0.80)'$, $vech(B) = (1.00, 0.00, 1.00)'$ (on the left), and (2) $vec(A) = (0.85, 0.20, 0.10, 0.80)'$, $vech(B) = (0.60, 0.30, 0.40)'$ (on the right). For each (quarterly) horizon, we compare (i) the aggregated Impulse Response Functions of both the mixed frequency procedures, and (ii) the quarterly IRFs (*naive* approach), relative to the monthly IRFs. Values near to one, indicate that the mixed frequency (or the quarterly frequency) impulse responses are close to the monthly impulse responses. In both the fourth and the eighth column, the value "NaN" is due to the Cholesky restriction.

2.4 Sum sampling

Considering the *sum-over-the low frequency period* sampling of the high frequency process in Eq. (2.2), we obtain a quite different low frequency solution *w.r.t.* that obtained in the point-in-time case. In particular, in the case of sum sampling, we multiply both sides of Eq. (2.2) by $\omega(L) = (I+L+L^2)$ and $\Gamma(L) = (I+AL+A^2L^2)$. The low frequency result is given by

$$\begin{aligned}
\Gamma(L)(I-AL)\omega(L)Y_t^* =& \Gamma(L)B\omega(L)\varepsilon_t \\
(I+AL+A^2L^2)(I-AL)(I+L+L^2)Y_t^* =& (I+AL+A^2L^2)B(I+L+L^2)\varepsilon_t \\
(I+L+L^2-A^3L^3-A^3L^4-A^3L^5)Y_t^* =& [I+(I+A)L+(I+A+A^2)L^2+ \\
& +(A^2+A)L^3+A^2L^4]B\varepsilon_t
\end{aligned}$$

hence,

$$\begin{aligned}
Y_t^* + Y_{t-1}^* + Y_{t-2}^* = A^3(Y_{t-3}^* + Y_{t-4}^* + Y_{t-5}^*)+ \\
+ B\varepsilon_t + (I+A)B\varepsilon_{t-1} + (I+A+A^2)B\varepsilon_{t-2}+ \\
+ (A+A^2)B\varepsilon_{t-3} + A^2B\varepsilon_{t-4}. \quad (2.22)
\end{aligned}$$

Since $Y_t^*+Y_{t-1}^*+Y_{t-2}^* = Y_\tau$ and $Y_{t-3}^*+Y_{t-4}^*+Y_{t-5}^* = Y_{\tau-1}$, the quarterly aggregated solution is defined by

$$Y_\tau = CY_{\tau-1} + \eta_\tau,$$

with η_τ the residual term of the quarterly process and $C = A^3$. However, summing the monthly values during each reference quarter, we can note that the residual component η_τ can be decomposed into the sum of two distinct elements:

$$\lambda_{\{t,t-1,t-2\}} = B\varepsilon_t + (I+A)B\varepsilon_{t-1} + (I+A+A^2)B\varepsilon_{t-2},$$

and

$$\lambda_{\{t-3,t-4,t-5\}} = (A+A^2)B\varepsilon_{t-3} + A^2B\varepsilon_{t-4}.$$

At quarterly frequency, $\lambda_{\{t,t-1,t-2\}}$ corresponds to the residual component obtained at time τ, and $\lambda_{\{t-3,t-4,t-5\}}$ to the sum of the residuals related to $\tau-1$. Then, since $\eta_\tau = (\lambda_{\{t,t-1,t-2\}} + \lambda_{\{t-3,t-4,t-5\}})$, we can note that

$$\mathbb{E}(\eta_\tau\eta'_{\tau-l}) = \begin{cases} \Omega & l=0 \\ \Phi(l)\equiv\Phi & l=\pm1 \\ \mathbf{0}_n & |l|\geq 2 \end{cases},$$

where $\Omega, \Phi \neq \mathbf{0}_n$, are respectively the covariance matrix and the first order autocovariance matrix of the residual η_τ. Specifically, η_τ represents the VMA(1) component of the aggregated solution. Rewriting η_τ as a VMA(1) process, i.e. $\eta_\tau = \xi_\tau + Q\xi_{\tau-1}$, with

$$\xi_\tau = \lambda_{\{t,t-1,t-2\}} = B\varepsilon_t + (I+A)B\varepsilon_{t-1} + (I+A+A^2)B\varepsilon_{t-2},$$

$$Q\xi_{\tau-1} = \lambda_{\{t-3,t-4,t-5\}} = (A+A^2)B\varepsilon_{t-3} + A^2B\varepsilon_{t-4},$$

the aggregated solution can be finally rewritten as the VARMA(1,1) process defined by

$$Y_\tau = CY_{\tau-1} + \xi_\tau + Q\xi_{\tau-1}, \qquad \xi_\tau \sim (0,\Pi) \tag{2.23}$$

with

$$\begin{aligned}\Pi =&\mathbb{E}(\xi_\tau\xi_\tau') = \\ =&BB' + (I+A)BB'(I+A)' + (I+A+A^2)BB'(I+A+A^2)',\end{aligned}$$

$$\begin{aligned}\Omega =&\mathbb{E}(\eta_\tau\eta_\tau') = \mathbb{E}((\xi_\tau + Q\xi_{\tau-1})(\xi_\tau + Q\xi_{\tau-1})') \\ =&BB' + (I+A)BB'(I+A)' + (I+A+A^2)BB'(I+A+A^2)' + \\ &\qquad + (A+A^2)BB'(A+A^2)' + A^2BB'A'^2 \qquad (2.24)\\ =&\Pi + Q\Pi Q' \qquad (2.25)\end{aligned}$$

$$\begin{aligned}\Phi =&\mathbb{E}[(\eta_\tau)(\eta_{\tau-1})'] = \mathbb{E}[(\xi_\tau + Q\xi_{\tau-1})(\xi_{\tau-1} + Q\xi_{\tau-2})'] = \\ =&\mathbb{E}[Q\xi_{\tau-1}\xi_{\tau-1}'] \\ =&\mathbb{E}\left\{[(A+A^2)B\varepsilon_{t-3} + A^2B\varepsilon_{t-4}][B\varepsilon_{t-3} + (I+A)B\varepsilon_{t-4} + (I+A+A^2)B\varepsilon_{t-5})]'\right\} \\ =&\mathbb{E}[(A+A^2)B\varepsilon_{t-3}\varepsilon_{t-3}'B'] + \mathbb{E}[A^2B\varepsilon_{t-4}\varepsilon_{t-4}'B'(I+A)'] \\ =&(A+A^2)BB' + A^2BB'(I+A)' \qquad (2.26)\\ =&Q\Pi \qquad (2.27)\end{aligned}$$

and $\Phi(j) = 0, \forall j \geq 2$.

2.4.1 Estimation approach

The two step procedure introduced in subsection 2.3.1 can be easily generalized to obtain the estimates of $\hat{A}$ and $\hat{B}$ of the monthly SVAR(1) process, from the estimated quarterly matrices $\hat{C}$, $\hat{\Omega}$ of Eq. (2.23), (2.24). Since the residual term $\eta_\tau \sim VMA(1)$, in this discussion we can't consider least square estimation of C and Ω. By this way we introduce the Instrumental variable estimator of the quarterly matrices.
In the general case, consider the model

$$Y_\tau = CX_\tau + \eta_\tau,$$

with $X_\tau = Y_{\tau-1}$, and C the $k \times 1$ parameter vector. Assume that exists a l-dimensional process $\{z_t\}$ of instruments with matrix $T \times l$ of observations $Z_\tau = (z'_\tau, z'_{\tau-1}, \ldots, z'_{\tau-l})'$, with l greater then the number of parameters in C, i.e. $l \geq k$. The process $\{z_t\}$ verify the following conditions:

$$\begin{aligned}
&\mathbb{E}(\eta_\tau) = 0, \\
&\mathbb{E}(\eta_\tau, \eta_{\tau-1}) = c_1, \\
&\mathbb{E}(\eta_\tau | z_\tau, z_{\tau-1}, \ldots, \eta_{\tau-2}, \eta_{\tau-3}, \ldots) = 0, \\
&W = \lim_{T\to\infty} (1/T)\mathbb{E}(Z'\eta_\tau\eta'_\tau Z) \qquad \text{exists and it is of full rank.}
\end{aligned}$$

Then, given the compact form $Y = XC' + \eta$, the Instrumental Variables estimator is given by

$$\hat{C}_{IV} = (X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'Y). \tag{2.28}$$

In this discussion we consider the instrumental variable estimator of the VAR coefficients matrix C and of the residual covariance matrix Ω. In particular, following Friedlander, Stoica and Söderström (1985) and Andrews (1991), we define the optimal choice of the matrix of instruments Z, represented by two lagged values of Y_τ, i.e. $Z_\tau = (Y'_{\tau-2}, Y'_{\tau-3})'$, we estimate the matrix of the VAR coefficients $\hat{C}$ as in equation Eq. (2.28) and, given the instrumental residuals

$$\hat{\eta}_{IV} = Y - X\hat{C}'_{IV},$$

we calculate the residual sample covariance matrix Ω. As in the case of the point-in-time solution, the idea is to obtain $\tilde{A}$ from the real cube root of $\hat{C}$, and then obtain $\hat{A}$ and $\hat{B}$ through the minimum distance approach, with starting values given by $\tilde{A}$. The *"low frequency - high frequency"* mapping is defined by

$$\begin{aligned}
&C = A^3 \\
&\Omega = BB' + (I + A)BB'(I + A)' + (I + A + A^2)BB'(I + A + A^2)' + \\
&\qquad\qquad + (A + A^2)BB'(A + A^2)' + A^2BB'A'^2.
\end{aligned}$$

As in Section 2.2, considering B non-recursive, the monthly SVAR(1) is not identified. We can summarize the correspondences between the number of free parameters of the quarterly and the monthly process by the following table:

	quarterly	monthly B triangular		monthly B non-triangular	
C	n^2	n^2	A	n^2	A
Ω	$n(n+1)/2$	$n(n+1)/2$	B	n^2	B
	$n^2 + \frac{n(n+1)}{2}$	$=$ $n^2 + \frac{n(n+1)}{2}$		$<$ $2n^2$	

Assuming B with N^2 parameters, in those cases in which we observe a structural change in Ω, we can solve the occurred identification problem as in Lanne and Lutkepohl (2008)[18]: (i) we estimate the VAR coefficient matrix C, (ii) we calculate the covariance matrices $\mathbb{E}(\eta_\tau, \eta_\tau) = \Omega_1$, $\tau = 3, \ldots, T_b$ and $\mathbb{E}(\eta_\tau, \eta_\tau) = \Omega_2$, $\tau = T_b + 3, \ldots, T$, with $V \neq I$, diagonal matrix. In these cases, the *"low frequency - high frequency"* mapping is represented by

$$
\begin{aligned}
&C = A^3,\\
&\Omega_1 = BB' + (I+A)BB'(I+A)' + (I+A+A^2)BB'(I+A+A^2)' +\\
&\qquad\qquad + (A+A^2)BB'(A+A^2)' + A^2BB'A'^2,\\
&\Omega_2 = BVB' + (I+A)BVB'(I+A)' + (I+A+A^2)BVB'(I+A+A^2)' +\\
&\qquad\qquad + (A+A^2)BVB'(A+A^2)' + A^2BVB'A'^2,
\end{aligned}
$$

implying exact identification, i.e.

	quarterly	monthly B non-triangular	
C	n^2	n^2	A
Ω_1	$n(n+1)/2$	n^2	B
Ω_2	$n(n+1)/2$	n	V
	$n^2 + n(n+1)$	$=$ $n^2 + n^2 + n$	

An interesting consideration concerns the estimation of the aggregated solution. With VARMA models the econometrician classically requires to impose some set of conditions to achieve uniqueness of the VARMA representation, i.e. echelon

[18] An alternative solution could be obtained by involving the estimated autocovariance matrix $\Phi(1)$.

form or final equation form[19]. In the specific case of aggregation sampling, after the estimation of the high frequency parameter matrices A and B, we are also able to identify all the quarterly parameters. In particular, once $\hat{A}$ and $\hat{B}$ are obtained, we can estimate indirectly the covariance matrix of ξ_τ, defined by

$$\hat{\Pi} = \hat{B}\hat{B}' + (I + \hat{A})\hat{B}\hat{B}'(I + \hat{A})' + (I + \hat{A} + \hat{A}^2)\hat{B}\hat{B}'(I + \hat{A} + \hat{A}^2)'.$$

Since Π is invertible and the autocovariance matrix Φ of the residual component η_τ can be rewritten as in Eq. (2.27), the VMA coefficient matrix Q can be obtained as

$$\hat{Q} = \hat{\Phi}\hat{\Pi}^{-1}, \tag{2.29}$$

with Φ obtainable either with indirect estimation (given the relation in Eq. (2.26)) or with the Instrumental Variable approach (see Cumby and Huizinga (1992) and Hall (1995)). In APPENDIX A4 we explain and provide further details about the indirect estimation of VMA matrix Q.

Minimum Distance estimation

In the case of sum sampling, the monthly unknown parameter vector $\theta = (vec(A)', vech(B)')'$ is function of the quarterly $n^2 + n(n+1)/2 \times 1$ vector $\hat{\phi} = (vec(\hat{C})', vech(\hat{\Omega})')'$. Without the inclusion of further restrictions, the order conditions are guaranteed, i.e.

$$dim(\hat{\phi}) = n^2 + n(n+1)/2 = n^2 + n(n+1)/2 = (vec(A)', vech(B)')' = dim(\theta).$$

The function h of Eq. (2.14), is summarized by the mapping

$$\begin{aligned} C &= A^3, \\ \Omega &= BB' + (I + A)BB'(I + A)' + (I + A + A^2)BB'(I + A + A^2)' + \\ &\quad + (A + A^2)BB'(A + A^2)' + A^2BB'A'^2, \end{aligned}$$

with weighting matrix $S = \hat{\Psi}^{-1}$ and $\hat{\Psi}$ built as the block-diagonal matrix given by

$$\hat{\Psi} = \begin{pmatrix} \hat{\Psi}_{\hat{C}} & 0 \\ 0 & \hat{\Psi}_{\hat{\Omega}} \end{pmatrix}.$$

In particular, $\hat{\Psi}_{\hat{C}}$ is the asymptotic covariance matrices of the quarterly estimates $\hat{C}$, and $\hat{\Psi}_{\hat{\Omega}}$ is the asymptotic covariance matrix of the quarterly estimates $\hat{\Omega}$. Following Cumby and Huizinga (1992), Newey and West (1994) and Hall (1995), we

[19]See Lutkepohl(2012).

obtain the expressions of asymptotic covariance matrices $\hat{\Psi}_{\hat{C}}$ and $\hat{\Psi}_{\hat{\Omega}}$ with the Instrumental Variables estimation approach[20].
In particular, given $X = (Y_0, Y_1, \ldots, Y_{T-1})$, $Y = (Y_1, Y_2, \ldots, Y_T)$ and the $h \times T$ matrix of instruments Z, we introduce the Instrumental Variables estimator $\hat{C} = (XZ'(Z'Z)^{-1}ZX')^{-1}XZ'(Z'Z)^{-1}Y$.

Likewise, considering B nontriangular, the monthly unknown parameter vector $\theta = (vec(A)', vec(B)', diag(V)')'$ is function of the quarterly $n^2 + n(n+1) \times 1$ vector $\hat{\phi} = (vec(\hat{C})', vech(\hat{\Omega}_1)', vech(\hat{\Omega}_2)')'$. The validity of order conditions can be summarized as

$$dim(\hat{\phi}) = n^2 + n(n+1) < 2n^2 + n = (vec(A)', vec(B)', diag(V)')' = dim(\theta).$$

The function h of Eq. (2.14), is summarized by the mapping

$$\begin{aligned} C &= A^3, \\ \Omega_1 &= BB' + (I+A)BB'(I+A)' + (I+A+A^2)BB'(I+A+A^2)' + \\ &\quad + (A+A^2)BB'(A+A^2)' + A^2BB'A'^2, \\ \Omega_2 &= BVB' + (I+A)BVB'(I+A)' + (I+A+A^2)BVB'(I+A+A^2)' + \\ &\quad + (A+A^2)BVB'(A+A^2)' + A^2BVB'A'^2, \end{aligned}$$

with weighting matrix $S = \hat{\Psi}^{-1}$ and $\hat{\Psi}$ built as the block-diagonal matrix given by

$$\hat{\Psi} = \begin{pmatrix} \hat{\Psi}_{\hat{C}} & 0 & 0 \\ 0 & \hat{\Psi}_{\hat{\Omega}1} & 0 \\ 0 & 0 & \hat{\Psi}_{\hat{\Omega}2} \end{pmatrix},$$

with $\hat{\Psi}_{\hat{\Omega}1}$ and $\hat{\Psi}_{\hat{\Omega}2}$ the asymptotic covariance matrices of the quarterly estimates $\hat{\Omega}_1$, $\hat{\Omega}_2$.

2.4.2 Monte Carlo experiments

We consider two different Monte Carlo experiments, with different sample sizes, to evaluate the solutions proposed above either for B lower triangular, or for B nontriangular, for the "sum" scheme.
For $M = 1000$ replications, we generate the monthly trivariate SVAR(1) process defined in Eq. (2.2), where Y_t^*, $t = 1\ldots, T$, is the $n \times 1$ vector of the n monthly series, Y_0 is set to $0_{n\times1}$, A is the $n \times n$ coefficient matrix, B is the $n \times n$ matrix of

[20] The alternative approach for the estimation of the asymptotic covariance matrices of the estimates is represented by the bootstrap techniques.

coefficients of instantaneous shocks. We consider different sample sizes, in particular $T = \{600, 1200, 3600\}$ (months, corresponding to $200, 400$ and 1200 quarters). For each replication, we apply the filters $\omega(L) = (I + L + L^2)$ and $\Gamma(L) = (I + AL + A^2L^2)$, to the monthly series Y_t^*. Hence we sum the monthly observations appearing during each reference quarter, obtaining the quarterly variable Y_τ, aggregated *via* sum sampling. The quarterly aggregated process is the VARMA(1,1) reported in Eq. (2.23).
In the case of B lower triangular, the estimation procedure, corresponds to:

- obtain the quarterly Instrumental Variables estimates of $\hat{C}$ and $\hat{\Omega}$, with the vector of instruments given by $Z_\tau = (Y'_{\tau-2}, Y'_{\tau-3})'$[21];
- solve the (real) cube root $\hat{C}^{1/3} = \tilde{A}$;
- use $\tilde{A}$ as starting values for the Minimum distance estimation of A and $\Sigma = BB'$, given the mapping of Eq. (2.16);
- use the Cholesky decomposition of $\hat{\Sigma}$ to obtain $\hat{B}$.

For the case of B non recursive, we refer to the monthly process in Eq. (2.9), where $V = diag(v_1, v_2, v_3)$, T_b is the break in the process and, for simplicity, we assume that (i) T_b coincides with the third month of the reference quarter, and (ii) $T_b = T/2$. In this case the estimation approach can be summarized by the following steps:

- estimate the quarterly matrices $\hat{C}$, $\hat{\Omega}_i$, with $i = 1, 2$;
- solve the cube root $\tilde{A} = \hat{C}^{1/3}$;
- estimate the monthly vector of parameters $(vec(\hat{A})', vech(\hat{\Omega}_1)', vech(\hat{\Omega}_2)')'$, solving the minimization problem in Eq. (2.14), with the mapping described by the system in Eq. (2.16);
- we consider a further Minimum distance step in which we estimate $\hat{B}$ and $\hat{V}$ from the relationships in Eq. (2.17).

As pointed out before, for the quarterly VARMA(1,1) process, we can also recover the quarterly parameters of Q. In particular, with $\hat{A}$ and $\hat{B}$, we can obtain an estimates of $var(\xi_\tau) = \hat{\Pi}$ and $cov(\eta_\tau, \eta_{\tau-1}) = \Phi$. Hence, since Π is non-singular, following the relationship in Eq. (2.29), we can obtain

$$\hat{Q} = \hat{\Phi}\hat{\Pi}^{-1}.$$

[21]See Amemiya (1974, 1977) and Friedlander, Stoica and Söderström (1985).

We report below two representative examples: a first design in which we consider B lower triangular ('Cholesky based'), and a second experiments with the structural matrix B leading for non-recursive schemes.
Referring to the first simulation design, we consider the monthly DGP

$$Y_t^* = \underbrace{\begin{pmatrix} 0.650 & 0.250 & 0.150 \\ 0.200 & 0.325 & 0.125 \\ 0.250 & 0.225 & 0.550 \end{pmatrix}}_{A} Y_{t-1}^* + \underbrace{\begin{pmatrix} 0.850 & 0 & 0 \\ 0.012 & 0.750 & 0 \\ 0.050 & 0.17 & 0.950 \end{pmatrix}}_{B} \varepsilon_t, \tag{2.30}$$

with $\varepsilon_t \sim (0, I)$. The monthly SVAR(1) process in Eq. (2.30) generates the quarterly VARMA(1,1) process defined by:

$$Y_\tau = \underbrace{\begin{pmatrix} 0.439 & 0.265 & 0.227 \\ 0.219 & 0.148 & 0.134 \\ 0.368 & 0.253 & 0.287 \end{pmatrix}}_{C} Y_{\tau-1} + \xi_\tau + \underbrace{\begin{pmatrix} 0.203 & 0.081 & 0.030 \\ 0.055 & 0.113 & 0.032 \\ 0.059 & 0.075 & 0.176 \end{pmatrix}}_{Q} \xi_{\tau-1},$$

with

$$\xi_\tau \sim \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \underbrace{\begin{pmatrix} 6.573 & 1.871 & 2.829 \\ 1.871 & 3.246 & 2.306 \\ 2.829 & 2.306 & 7.539 \end{pmatrix}}_{\Sigma} \right),$$

$$\xi_\tau + Q\xi_{\tau-1} = \eta_\tau \sim \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \underbrace{\begin{pmatrix} 8.220 & 2.694 & 4.202 \\ 2.694 & 3.721 & 3.110 \\ 4.202 & 3.110 & 9.168 \end{pmatrix}}_{\Omega} \right),$$

and

$$\mathbb{E}(\eta_\tau, \eta_{\tau-1}) = \underbrace{\begin{pmatrix} 1.572 & 0.712 & 0.986 \\ 0.662 & 0.543 & 0.654 \\ 1.023 & 0.759 & 1.663 \end{pmatrix}}_{\Phi}.$$

The representative design of Eq. (2.30), the related Monte Carlo results and the Impulse Response Functions for the sample size T=3600, are reported in Table 2.8, Table 2.9 and Figure 2.5.
As in the case of the point-in-time sampling, we evaluate the proposed estimation procedure referring to two measures of bias, in particular $bias(1)$, obtained as

the sum of the differences, in absolute value, of the population parameters and the Monte Carlo estimates, and $bias(2)$, calculated as $bias(1)/$(number of estimated elements). As measures of variability we report the mean across replications of the standard errors (in the Tables as "s.e.") obtained as the square root of the diagonal entries of the estimated asymptotic covariance matrices of the Minimum distance estimates (see APPENDIX 3), and the Monte Carlo standard errors (in the Tables as "MCs.e.").
The results shown in Table 2.8 and Table 2.9 confirm the conclusions reported for the point-in-time solutions: increasing the sample size we can observe an enhancement in the performances of the proposed approach and a strong reduction of the standard errors of the estimates. The reliability of the Minimum distance IRFs with respect to the DGP counterpart can be noted in Figure 2.5. We depict the response of the i-th variable to a shock on the j-th variable as $(j \to i)$ i.e. $(impulse \to response)$. Graphically we can note that the responses obtained with the proposed approach (black-solid line) approximate the true IRFs (red-dotted line) in an accurate way.

	T_m=600			T_m=1200			T_m=3600		
$vec(A)$	$vec(\hat{A})$	MCs.e.	s.e.	$vec(\hat{A})$	MCs.e	s.e.	$vec(\hat{A})$	MCs.e.	s.e.
0.650	0.624	0.12	0.13	0.621	0.10	0.09	0.638	0.07	0.04
0.200	0.220	0.11	0.10	0.215	0.10	0.08	0.211	0.07	0.03
0.250	0.271	0.15	0.12	0.278	0.14	0.12	0.262	0.09	0.05
0.250	0.285	0.13	0.09	0.282	0.10	0.06	0.269	0.08	0.06
0.325	0.249	0.14	0.12	0.282	0.13	0.10	0.294	0.10	0.08
0.225	0.264	0.09	0.08	0.237	0.07	0.05	0.251	0.09	0.06
0.150	0.157	0.11	0.08	0.164	0.10	0.08	0.153	0.06	0.04
0.125	0.146	0.07	0.07	0.132	0.07	0.06	0.131	0.05	0.05
0.550	0.507	0.20	0.13	0.515	0.19	0.12	0.524	0.13	0.09
bias(1)	0.288	-	-	0.215	-	-	0.146	-	-
bias(2)	0.032	-	-	0.024	-	-	0.016	-	-
$vech(B)$	$vech(\hat{B})$	MCs.e.	s.e.	$vech(\hat{B})$	MCs.e.	s.e	$vech(\hat{B})$	MCs.e.	s.e
0.850	0.883	0.18	0.15	0.886	0.13	0.12	0.874	0.08	0.07
0.012	-0.009	0.04	0.03	-0.004	0.03	0.02	0.001	0.01	0.00
0.050	0.013	0.03	0.02	0.001	0.04	0.02	0.031	0.02	0.02
0.750	0.750	0.16	0.15	0.740	0.16	0.14	0.749	0.15	0.10
0.170	0.157	0.13	0.09	0.191	0.08	0.06	0.174	0.05	0.05
0.950	0.882	0.31	0.30	0.887	0.20	0.19	0.905	0.18	0.16
bias(1)	0.430	-	-	0.196	-	-	0.101	-	-
bias(2)	0.072	-	-	0.033	-	-	0.017	-	-

Table 2.8: Monte Carlo results obtained for *sum-over-the quater* sampling of a SVAR(1) with B Cholesky-based. In the first column we report the elements of the DGP matrices A and B (of Eq. (2.30)). We consider $T = 600, 1200, 3600$, and we evaluate the results with two measures of bias: $bias(1)$ is calculated as the sum of the differences in absolute value of the "true" value of the elements and the Monte Carlo estimates; $bias(2)$ is calculated as $bias(1)$/(number of elements).

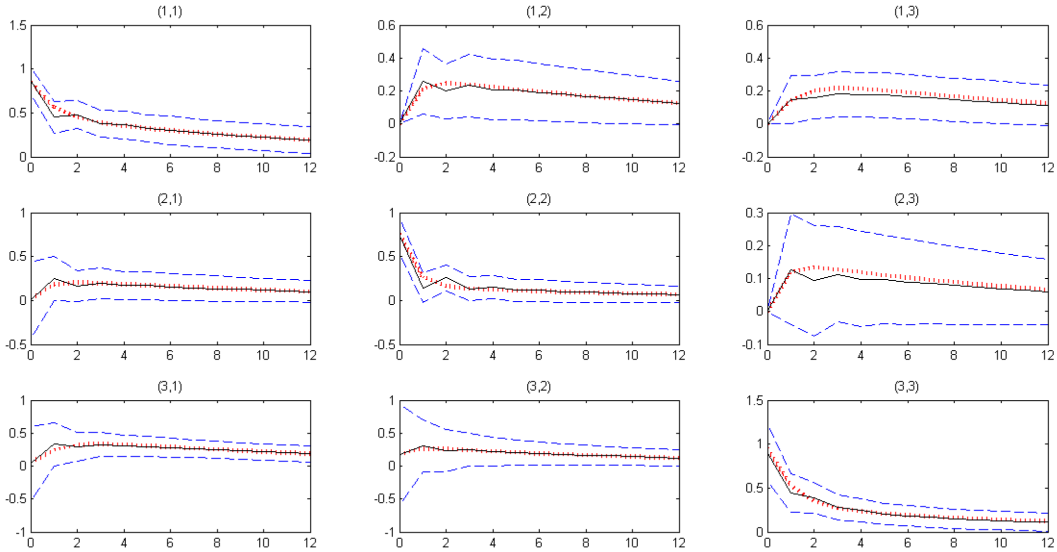


Figure 2.5: Impulse responses from Monte Carlo experiment reported in Table 2.8, with sample size $T = 3600$. The red-dotted line corresponds to the population responses, the blue-dashed lines refer to the confidence bounds (calculated as two-standar error bounds) of the IRFs calculated with the proposed Minimum distance approach (black line)

.

	T_m=600		T_m=1200		T_m=3600	
$vec(Q)$	$vec(\hat{Q})$	MCs.e.	$vec(\hat{Q})$	MCs.e	$vec(\hat{Q})$	MCs.e.
0.203	0.188	0.082	0.186	0.082	0.194	0.063
0.055	0.060	0.066	0.058	0.061	0.056	0.056
0.059	0.065	0.082	0.069	0.082	0.067	0.069
0.081	0.091	0.132	0.090	0.129	0.086	0.109
0.113	0.090	0.136	0.103	0.127	0.109	0.114
0.075	0.090	0.156	0.084	0.153	0.085	0.136
0.030	0.031	0.071	0.036	0.070	0.034	0.057
0.032	0.039	0.065	0.035	0.064	0.033	0.051
0.176	0.152	0.098	0.154	0.098	0.158	0.084
bias(1)	0.080	-	0.058	-	0.042	-
bias(2)	0.009	-	0.006	-	0.005	-

Table 2.9: Monte Carlo results for the estimation of the VMA coefficient matrix Q. The VARMA(1,1) is obtained from *sum-over-the quater* sampling of the SVAR(1) with B Cholesky-based, reported in Eq. (2.30). In the first column we report the DGP elements of Q. We consider $T = 600, 1200, 3600$, and we evaluate the results with two measures of bias: $bias(1)$ is calculated as the sum of the differences in absolute value of the "true" value of the elements and the Monte Carlo estimates; $bias(2)$ is calculated as $bias(1)$/(number of elements).

A second exercise is considered with B non-recursive. In particular we specify the monthly Monte Carlo design given by

$$Y_t^* = \underbrace{\begin{pmatrix} 0.850 & -0.100 & 0.200 \\ -0.100 & 0.400 & 0.650 \\ 0.150 & 0.500 & 0.300 \end{pmatrix}}_{A} Y_{t-1}^* + \underbrace{\begin{pmatrix} 0.750 & 0.400 & 0.300 \\ 0.500 & 0.850 & -0.100 \\ -0.500 & -0.150 & 0.500 \end{pmatrix}}_{B} \varepsilon_t, \quad (2.31)$$

with $\varepsilon_t \sim (0, I)$ for $t = 1, \ldots, T_b$ and $\varepsilon_t \sim (0, V)$ for $t = T_b + 1, \ldots, T$ and $V = diag(0.50, 1.40, 2.20)$.
The monthly SVAR(1) process in Eq. (2.31) generates the quarterly VARMA(1,1) process defined by:

$$Y_\tau = \underbrace{\begin{pmatrix} 0.676 & -0.004 & 0.186 \\ -0.076 & 0.418 & 0.447 \\ 0.137 & 0.344 & 0.376 \end{pmatrix}}_{C} Y_{\tau-1} + \xi_\tau + Q_i \xi_{\tau-1},$$

with

$$Q_1 = \begin{pmatrix} 0.216 & 0.137 & -0.172 \\ -0.028 & 0.163 & 0.136 \\ 0.009 & 0.229 & -0.018 \end{pmatrix}, \qquad Q_2 = \begin{pmatrix} -0.096 & 0.228 & 0.245 \\ -0.167 & 0.251 & 0.219 \\ -0.171 & 0.313 & 0.154 \end{pmatrix},$$

For the first part of the sample ($\tau = 3, \ldots, T_b$) we consider

$$\xi_\tau \sim \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \underbrace{\begin{pmatrix} 7.855 & 3.891 & 0.884 \\ 3.891 & 4.351 & 0.866 \\ 0.884 & 0.866 & 2.237 \end{pmatrix}}_{\Sigma} \right),$$

$$\xi_\tau + Q\xi_{\tau-1} = \eta_\tau \sim \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \underbrace{\begin{pmatrix} 9.873 & 4.397 & 1.876 \\ 4.397 & 5.003 & 1.488 \\ 1.876 & 1.488 & 3.059 \end{pmatrix}}_{\Omega} \right), \qquad \Phi = \underbrace{\begin{pmatrix} 2.081 & 1.290 & -0.075 \\ 0.533 & 0.718 & 0.421 \\ 0.944 & 1.016 & 0.167 \end{pmatrix}}_{\Phi}.$$

With $\tau = T_b + 3, \ldots, T$, the quarterly matrices Σ, Ω and Φ are defined as

$$\Sigma = \begin{pmatrix} 7.659 & 4.603 & 3.427 \\ 4.603 & 5.805 & 1.844 \\ 3.427 & 1.844 & 3.915 \end{pmatrix},$$

$$\Omega = \begin{pmatrix} 9.891 & 5.738 & 4.842 \\ 5.738 & 6.843 & 2.908 \\ 4.842 & 2.908 & 5.161 \end{pmatrix}, \qquad \Phi = \begin{pmatrix} 1.155 & 1.334 & 1.051 \\ 0.626 & 1.092 & 0.747 \\ 0.664 & 1.318 & 0.597 \end{pmatrix}.$$

The representative design of Eq. (2.31), the related Monte Carlo results and the Impulse Responses Functions for the sample size $T = 3600$, are respectively reported in Table 2.10, Table 2.11 and Figure 2.6 and Figure 2.7.
From the results in Table 2.10 we can note a non-negligible enhancement in the estimates of $\hat{A}$ and $\hat{B}$, increasing the sample size. In particular, both the measures of estimation bias and the standard errors of the estimates significantly diminish: since the estimates of the quarterly coefficients are more close to the DGP values when T becomes large, more accurate are the monthly estimates and then the impulse responses.
In Table 2.11 we show the estimates of the VMA(1) coefficients matrix, for the two periods, Q_1 for $\tau = 3, 6, \ldots, T_b$ and Q_2 for $\tau = T_b + 3, T_b + 6, \ldots, T$. The measures of distortion associated to the indirect estimates of $\hat{Q}$, seems to confirm the results obtained at the first step, likewise the impulse responses in Figure 2.6 and Figure 2.7.

	T_m=600			T_m=1200			T_m=3600		
$vec(A)$	$vec(\hat{A})$	MCs.e.	s.e.	$vec(\hat{A})$	MCs.e	s.e.	$vec(\hat{A})$	MCs.e.	s.e.
0.850	0.893	0.15	0.14	0.887	0.14	0.11	0.872	0.12	0.09
-0.100	-0.099	0.07	0.06	-0.098	0.06	0.04	-0.098	0.05	0.05
0.150	0.169	0.07	0.06	0.168	0.06	0.04	0.157	0.05	0.04
-0.100	-0.002	0.11	0.12	-0.019	0.11	0.12	-0.041	0.09	0.06
0.400	0.434	0.17	0.15	0.424	0.15	0.13	0.423	0.11	0.11
0.500	0.561	0.19	0.14	0.564	0.18	0.11	0.523	0.17	0.15
0.200	0.063	0.14	0.15	0.085	0.14	0.16	0.120	0.13	0.18
0.650	0.618	0.21	0.19	0.629	0.19	0.16	0.627	0.15	0.13
0.300	0.227	0.23	0.18	0.223	0.22	0.14	0.272	0.17	0.14
bias(1)	0.498	-	-	0.440	-	-	0.267	-	-
bias(2)	0.055	-	-	0.049	-	-	0.030	-	-
$vec(B)$	$vec(\hat{B})$	MCs.e.	s.e.	$vec(\hat{B})$	MCs.e.	s.e	$vec(\hat{B})$	MCs.e.	s.e
0.750	0.751	0.05	0.07	0.753	0.02	0.04	0.750	0.01	0.02
0.500	0.486	0.13	0.13	0.495	0.08	0.06	0.498	0.04	0.04
-0.500	-0.500	0.14	0.10	-0.505	0.09	0.06	-0.499	0.05	0.03
0.400	0.408	0.05	0.12	0.403	0.02	0.07	0.400	0.01	0.04
0.850	0.820	0.11	0.09	0.830	0.10	0.05	0.845	0.09	0.03
-0.150	-0.141	0.06	0.19	-0.145	0.06	0.12	-0.148	0.05	0.07
0.300	0.295	0.04	0.07	0.299	0.01	0.05	0.300	0.01	0.03
-0.100	-0.096	0.07	0.16	-0.094	0.05	0.10	-0.100	0.04	0.06
0.500	0.482	0.11	0.06	0.493	0.10	0.04	0.499	0.08	0.02
bias(1)	0.090	-	-	0.054	-	-	0.011	-	-
bias(2)	0.010	-	-	0.006	-	-	0.001	-	-
$diag(V)$	$diag(\hat{V})$	MCs.e.	s.e.	$diag(\hat{V})$	MCs.e.	s.e	$diag(\hat{V})$	MCs.e.	s.e
0.500	0.495	0.06	0.08	0.495	0.04	0.04	0.500	0.02	0.02
1.400	1.401	0.10	0.23	1.395	0.05	0.13	1.399	0.02	0.08
2.200	2.198	0.03	0.53	2.199	0.02	0.32	2.200	0.01	0.17
bias(1)	0.007	-	-	0.012	-	-	0.001	-	-
bias(2)	0.002	-	-	0.004	-	-	0.000	-	-

Table 2.10: Monte Carlo results obtained for *sum-over-the quater* sampling of a SVAR(1) with B non-recursive. In the first column we report the elements of the DGP matrices A and B (of Eq. (2.31)). We consider $T = 600, 1200, 3600$, and we evaluate the results with two measures of bias: $bias(1)$ is calculated as the sum of the differences in absolute value of the "true" value of the elements and the Monte Carlo estimates; $bias(2)$ is calculated as $bias(1)$/(number of elements).

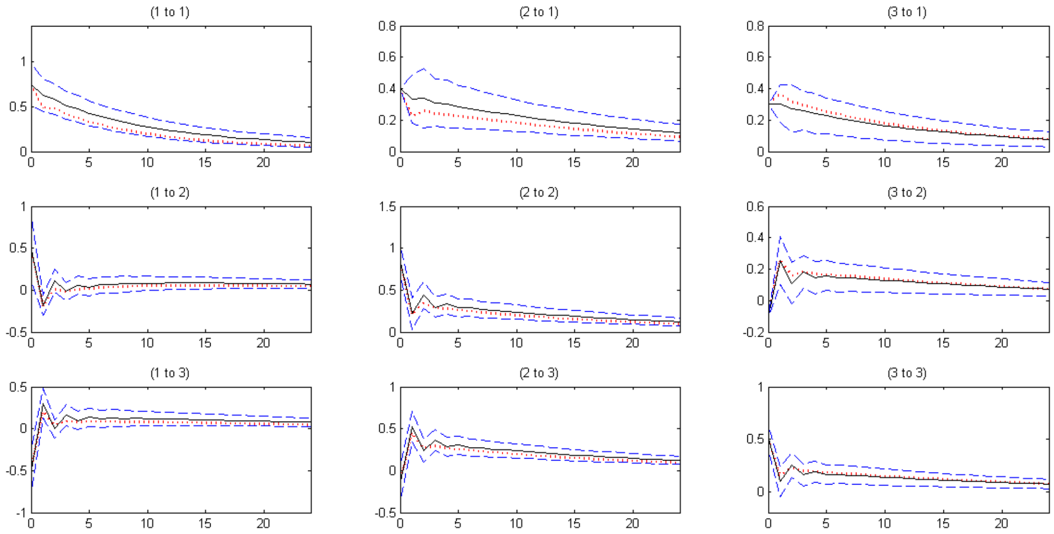
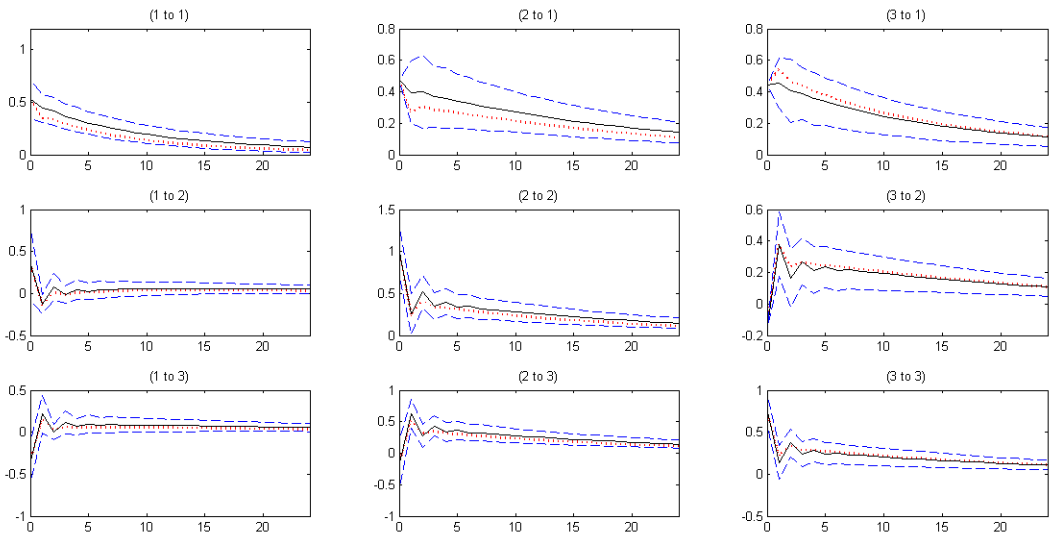


Figure 2.6: Impulse responses from Monte Carlo experiment reported in Table 2.10, with sample size $T = 3600$ and $t = 1, \ldots, T_b$. The red-dotted line corresponds to the population responses, the blue-dashed lines refer to the confidence bounds (calculated as two-standar error bounds) of the IRFs calculated with the proposed Minimum distance approach (black line)



.

Figure 2.7: Impulse responses from Monte Carlo experiment reported in Table 2.10, with sample size $T = 3600$ and $t = T_b + 1, \ldots, T$. The red-dotted line corresponds to the population responses, the blue-dashed lines refer to the confidence bounds (calculated as two-standar error bounds) of the IRFs calculated with the proposed Minimum distance approach (black line).

	T_m=600		T_m=1200		T_m=3600	
$vec(Q_1)$	$vec(\hat{Q}_1)$	MCs.e.	$vec(\hat{Q}_1)$	MCs.e	$vec(\hat{Q}_1)$	MCs.e.
0.216	0.180	0.08	0.193	0.05	0.203	0.04
-0.028	-0.023	0.03	-0.021	0.02	-0.021	0.02
0.009	-0.016	0.05	-0.010	0.03	0.001	0.02
0.137	0.280	0.17	0.254	0.14	0.207	0.13
0.163	0.204	0.08	0.201	0.06	0.190	0.07
0.229	0.278	0.12	0.270	0.11	0.234	0.10
-0.172	-0.361	0.24	-0.322	0.16	-0.257	0.15
0.136	0.070	0.18	0.080	0.09	0.093	0.09
-0.018	-0.086	0.15	-0.078	0.14	-0.036	0.14
bias(1)	0.623	-	0.510	-	0.274	-
bias(2)	0.069	-	0.057	-	0.030	-
$vec(Q_2)$	$vec(\hat{Q}_2)$	MCs.e.	$vec(\hat{Q}_2)$	MCs.e	$vec(\hat{Q}_2)$	MCs.e.
-0.096	-0.077	0.13	-0.065	0.05	-0.068	0.04
-0.167	-0.109	0.07	-0.108	0.05	-0.110	0.04
-0.171	-0.220	0.17	-0.203	0.07	-0.196	0.05
0.228	0.344	0.17	0.312	0.10	0.277	0.09
0.251	0.269	0.07	0.264	0.06	0.253	0.05
0.313	0.393	0.17	0.371	0.10	0.338	0.07
0.245	0.113	0.28	0.123	0.12	0.164	0.11
0.219	0.138	0.09	0.148	0.06	0.159	0.07
0.154	0.101	0.17	0.105	0.10	0.138	0.10
bias(1)	0.608	-	0.519	-	0.342	-
bias(2)	0.068	-	0.058	-	0.038	-

Table 2.11: Monte Carlo results for the estimation of the VMA coefficient matrix Q_1 and Q_2. The VARMA(1,1) is obtained from *sum-over-the quater* sampling of the SVAR(1) with B non-recursive, reported in Eq. (2.31). In the first column we report the DGP elements of Q_i, $i = 1, 2$. We consider $T = 600, 1200, 3600$, and we evaluate the results with two measures of bias: $bias(1)$ is calculated as the sum of the differences in absolute value of the "true" value of the elements and the Monte Carlo estimates; $bias(2)$ is calculated as $bias(1)$/(number of elements).

2.4.3 Comparison with state space estimation

In this section we consider the four Monte Carlo experiments examined in section 2.3.3 for the point-in-time case. Also for the *sum* aggregation, we compare the results obtained with state space estimation and the proposed approach in term of parameters estimation and impulse responses. The general design corresponds to the high frequency bivariate VAR(1) DGP of Eq. (2.20) and Eq. (2.21).

Results

To make results as comparable as possible, we refer to the same DGPs matrices considered in section 2.3.3 and we maintain the same measures of evaluation of the results: we compare the estimates with the measure *bias(1)*, considered above, and we evaluate the impulse responses with the ratio of (i) the root of the mean across replications of the squared error calculated for the mixed frequency IRFs (either with state space, or minimum distance) and (ii) root of the mean across replications of the RMSEs of the monthly IRFs, calculated with ordinary least squares.
The results obtained from the Monte Carlo experiments in the case of the sum sampling scheme confirm the considerations proposed above for the point-in-time sampling: the ratios related to the Minimum Distance estimation are a little bit bigger then the ratios calculated with state space methodology. In the case of $B \neq I$, the enhancement highlighted in the previous sections for the responses after a shock of the low frequency variable, appears slightly reduced, in particular for case "y to y", the responses of the low frequency variable (the first column of Tables 2.13 and 2.14): specifically, most of the ratios calculated with the minimum distance estimates become bigger then the state space counterparts.
As in the point-in-time comparison, we report also the ratios obtained with the aggregated mixed frequency IRFs and the low frequency results calculated through the *naive* approach. Specifically, we calculate the impulse responses for the monthly, the mixed frequency (with both approaches) and the quarterly data. To make IRFs comparable, we aggregate the monthly and the mixed frequency results through the aggregation scheme of sum sampling. The ratios reported in Table 2.15 and Table 2.16 confirm the consideration about temporal aggregation: the ratios related to the quarterly impulses are very far from the unit, especially at the instant related to the impact of the shock ($h = 0$). In other words, in those cases in which we assume that the DGP is at high frequency, both the mixed frequency approaches alleviate the effects of temporal aggregation.

Design: vec(A)'=(0.75 , 0.20 , 0.10, 0.70)

vech(B)'=(1.00 , 0.00 , 1.00)			vech(B)'=(0.60 , 0.30 , 0.40)		
TRUE	St.Sp.	Min.Dist.	TRUE	St.Sp	Min.Dist.
$vec(A)$	$vec(\hat{A})$	$vec(\hat{A})$	$vec(A)$	$vec(\hat{A})$	$vec(\hat{A})$
0.75	0.74	0.76	0.75	0.73	0.76
	(0.06)	(0.03)		(0.10)	(0.05)
0.20	0.21	0.19	0.20	0.20	0.21
	(0.05)	(0.03)		(0.07)	(0.06)
0.10	0.11	0.09	0.10	0.12	0.09
	(0.06)	(0.04)		(0.09)	(0.07)
0.70	0.69	0.71	0.70	0.70	0.70
	(0.04)	(0.05)		(0.06)	(0.08)
bias(1)	0.03	0.04	bias(1)	0.01	0.03
$vech(BB')$	$vech(\hat{B}\hat{B}')$	$vech(\hat{B}\hat{B}')$	$vech(BB)$	$vech(\hat{B}\hat{B}')$	$vech(\hat{B}\hat{B}')$
1.00	1.02	0.97	0.36	0.37	0.35
	(0.19)	(0.17)		(0.07)	(0.07)
0.00	-0.01	0.01	0.18	0.18	0.17
	(0.12)	(0.05)		(0.04)	(0.03)
1.00	0.99	0.97	0.25	0.25	0.24
	(0.09)	(0.11)		(0.02)	(0.09)
bias(1)	0.04	0.07	bias(1)	0.01	0.03

Design: vec(A)'=(0.85 , 0.20 , 0.10, 0.80)

vech(B)'=(1.00 , 0.00 , 1.00)			vech(B)'=(0.60 , 0.30 , 0.40)		
$vec(A)$	$vec(\hat{A})$	$vec(\hat{A})$	$vec(A)$	$vec(\hat{A})$	$vec(\hat{A})$
0.85	0.84	0.86	0.85	0.84	0.86
	(0.05)	(0.04)		(0.08)	(0.07)
0.20	0.20	0.19	0.20	0.20	0.20
	(0.04)	(0.05)		(0.06)	(0.09)
0.10	0.10	0.08	0.10	0.10	0.08
	(0.04)	(0.06)		(0.07)	(0.11)
0.80	0.79	0.81	0.80	0.79	0.80
	(0.03)	(0.08)		(0.03)	(0.03)
bias(1)	0.02	0.05	bias(1)	0.03	0.04
$vech(BB')$	$vech(\hat{B}\hat{B}')$	$vech(\hat{B}\hat{B}')$	$vech(BB')$	$vech(\hat{B}\hat{B}')$	$vech(\hat{B}\hat{B}')$
1.00	1.00	0.96	0.36	0.36	0.35
	(0.17)	(0.16)		(0.06)	(0.11)
0.00	-0.01	0.03	0.18	0.18	0.18
	(0.11)	(0.07)		(0.03)	(0.04)
1.00	0.99	0.97	0.25	0.25	0.25
	(0.08)	(0.17)		(0.02)	(0.13)
bias(1)	0.02	0.10	bias(1)	0.01	0.02

Table 2.12: Estimates from two bivariate SVAR(1) data generating processes, with $T = 300$ high frequency instants. We report the DGP values in the first column, the estimates obtained with state space approach, and the minimum distance estimates (both with related empirical standard errors). In both the experiments of the first half part of the table $B = I$: on the left $vec(A) = (0.75, 0.20, 0.10, 0.70)'$ and on the right part $vec(A) = (0.85, 0.20, 0.10, 0.80)'$. In the second half part of the table $vech(B) = (0.60, 0.30, 0.40)'$, with $vec(A) = (0.75, 0.20, 0.10, 0.70)'$ on the left and $vec(A) = (0.85, 0.20, 0.10, 0.80)'$ on the right.

Design:	$vec(A) = (0.75, 0.20, 0.10, 0.70)'$, $vech(B) = (1.0, 0.0, 1.0)$							
	St.Sp vs monthly				Min.Dis vs monthly			
h (months)	y to y	x to y	y to x	x to x	y to y	x to y	y to x	x to x
0	2.27	2.13	NaN	1.14	2.73	2.68	NaN	2.64
1	1.25	1.39	1.35	1.11	1.48	1.63	1.65	1.28
2	1.19	1.16	1.30	1.05	1.05	1.22	1.44	1.10
3	1.15	1.09	1.26	1.04	1.07	1.17	1.40	1.10
4	1.12	1.07	1.24	1.05	1.09	1.13	1.38	1.07
5	1.08	1.05	1.21	1.07	1.11	1.11	1.38	1.10
6	1.06	1.03	1.20	1.09	1.13	1.10	1.37	1.15
7	1.04	1.02	1.18	1.10	1.16	1.11	1.36	1.19
8	1.03	1.02	1.17	1.11	1.18	1.13	1.37	1.23
9	1.02	1.01	1.16	1.12	1.22	1.15	1.39	1.27
10	1.02	1.01	1.15	1.12	1.24	1.18	1.40	1.30
11	1.02	1.01	1.14	1.12	1.27	1.25	1.42	1.35
12	1.02	1.00	1.14	1.11	1.30	1.27	1.45	1.39

Design:	$vec(A) = (0.75, 0.20, 0.10, 0.70)'$, $vech(B) = (0.6, 0.3, 0.4)$							
	St.Sp vs monthly				Min.Dis vs monthly			
h (months)	y to y	x to y	y to x	x to x	y to y	x to y	y to x	x to x
0	2.44	2.00	NaN	1.71	2.98	3.68	NaN	3.45
1	1.37	1.24	1.73	1.18	2.07	2.34	2.23	1.95
2	1.21	1.08	1.52	1.20	1.05	1.20	1.31	1.11
3	1.16	1.05	1.43	1.23	1.13	1.12	1.31	1.09
4	1.12	1.05	1.38	1.25	1.14	1.10	1.26	1.10
5	1.10	1.04	1.35	1.26	1.20	1.12	1.26	1.11
6	1.09	1.03	1.33	1.26	1.24	1.16	1.26	1.14
7	1.08	1.02	1.31	1.26	1.28	1.20	1.27	1.16
8	1.07	1.02	1.30	1.26	1.31	1.25	1.28	1.19
9	1.06	1.02	1.28	1.25	1.35	1.30	1.30	1.21
10	1.06	1.02	1.27	1.24	1.38	1.34	1.31	1.24
11	1.06	1.02	1.25	1.24	1.42	1.39	1.33	1.27
12	1.06	1.02	1.24	1.23	1.45	1.43	1.35	1.30

Table 2.13: In the Table we show the RMSE ratios calculated for the evaluation of the Impulse Response Functions for the two mixed frequency procedure. We compare (i) the RMSE for the high frequency IRFs obtained from state space and minimum distance estimation, with (ii) the (benchmark) RMSEs obtained from the estimates the high frequency parameters, once we assume that all the variables can be observed: more the rates are near to one, more the mixed frequency impulse responses are close to the monthly impulse responses. In both the fourth and the eighth column, the value "NaN" is due to the Cholesky restriction.

Design:	$vec(A) = (0.85, 0.20, 0.10, 0.80)'$, $vech(B) = (1.0, 0.0, 1.0)$							
	St.Sp vs monthly				Min.Dis vs monthly			
h (months)	y to y	x to y	y to x	x to x	y to y	x to y	y to x	x to x
0	2.14	1.98	NaN	1.10	2.31	2.41	NaN	2.20
1	1.29	1.46	1.34	1.14	1.33	1.55	1.51	1.34
2	1.17	1.22	1.31	1.11	1.14	1.27	1.49	1.14
3	1.15	1.13	1.28	1.09	1.12	1.20	1.48	1.10
4	1.14	1.10	1.26	1.09	1.12	1.18	1.47	1.09
5	1.12	1.08	1.24	1.10	1.11	1.16	1.44	1.07
6	1.11	1.07	1.23	1.11	1.10	1.13	1.42	1.07
7	1.09	1.06	1.21	1.12	1.10	1.11	1.40	1.08
8	1.08	1.06	1.20	1.13	1.08	1.09	1.37	1.09
9	1.07	1.05	1.18	1.13	1.07	1.09	1.35	1.11
10	1.06	1.05	1.17	1.13	1.06	1.07	1.32	1.12
11	1.05	1.04	1.16	1.13	1.05	1.07	1.30	1.13
12	1.04	1.04	1.15	1.13	1.04	1.06	1.27	1.14

Design:	$vec(A) = (0.85, 0.20, 0.10, 0.80)'$, $vech(B) = (0.6, 0.3, 0.4)$							
	St.Sp vs monthly				Min.Dis vs monthly			
h (months)	y to y	x to y	y to x	x to x	y to y	x to y	y to x	x to x
0	2.14	1.81	NaN	1.66	2.20	2.25	NaN	2.47
1	1.35	1.29	1.47	1.19	1.43	1.54	1.19	1.31
2	1.18	1.12	1.40	1.15	1.14	1.26	1.15	1.09
3	1.13	1.07	1.35	1.16	1.13	1.14	1.13	1.05
4	1.10	1.05	1.32	1.18	1.04	1.07	1.11	1.05
5	1.07	1.05	1.30	1.19	1.01	1.03	1.11	1.06
6	1.06	1.04	1.28	1.20	1.00	1.01	1.10	1.07
7	1.04	1.04	1.27	1.21	1.00	1.00	1.09	1.09
8	1.03	1.03	1.26	1.21	1.00	0.99	1.08	1.10
9	1.03	1.03	1.25	1.21	1.00	0.99	1.08	1.12
10	1.02	1.02	1.24	1.21	1.00	0.99	1.07	1.13
11	1.02	1.02	1.23	1.21	0.99	0.99	1.07	1.15
12	1.01	1.02	1.22	1.21	0.98	0.99	1.06	1.16

Table 2.14: In the Table we show the RMSE ratios calculated for the evaluation of the Impulse Response Functions for the two mixed frequency procedure. We compare (i) the RMSE for the high frequency IRFs obtained from state space and minimum distance estimation, with (ii) the (benchmark) RMSEs obtained from the estimates the high frequency parameters, once we assume that all the variables can be observed: more the rates are near to one, more the mixed frequency impulse responses are close to the monthly impulse responses. In both the fourth and the eighth column, the value "NaN" is due to the Cholesky restriction.

	$vec(A) = (0.75, 0.20, 0.10, 0.70)'$, $vech(B) = (1.00, 0.00, 1.00)'$				$vec(A) = (0.75, 0.20, 0.10, 0.70)'$, $vech(B = (0.60, 0.30, 0.40)'$			
				State Space vs monthly				
h (quarters)	y to y	x to y	y to x	x to x	y to y	x to y	y to x	x to x
0	1.26	1.54	NaN	1.10	1.29	1.35	NaN	1.26
1	1.13	1.05	1.27	1.06	1.11	1.05	1.39	1.25
2	1.06	1.04	1.21	1.13	1.06	1.03	1.32	1.26
3	1.04	1.03	1.17	1.15	1.05	1.03	1.26	1.25
4	1.05	1.03	1.15	1.15	1.05	1.03	1.22	1.22
5	1.06	1.04	1.14	1.14	1.06	1.04	1.18	1.19
6	1.07	1.05	1.15	1.15	1.06	1.05	1.16	1.16
7	1.09	1.07	1.16	1.15	1.07	1.06	1.14	1.14
8	1.11	1.09	1.17	1.16	1.08	1.07	1.12	1.12
				Minimum Distance vs monthly				
0	1.30	1.62	NaN	1.33	1.35	1.61	NaN	1.47
1	1.08	1.15	1.40	1.15	1.16	1.10	1.27	1.06
2	1.16	1.15	1.38	1.25	1.27	1.18	1.27	1.13
3	1.26	1.23	1.43	1.34	1.37	1.32	1.32	1.22
4	1.37	1.34	1.51	1.44	1.46	1.44	1.39	1.32
5	1.50	1.47	1.62	1.56	1.57	1.56	1.48	1.42
6	1.63	1.60	1.74	1.69	1.67	1.68	1.57	1.52
7	1.77	1.74	1.87	1.81	1.78	1.79	1.67	1.62
8	1.82	1.87	1.90	1.84	1.90	1.91	1.77	1.73
				quarterly vs monthly				
0	7.75	3.67	NaN	7.54	8.39	6.47	NaN	5.63
1	3.66	3.12	1.46	2.74	3.95	3.76	1.24	1.80
2	3.05	2.90	1.60	2.05	3.16	3.08	1.31	1.46
3	2.86	2.83	1.74	1.91	2.88	2.88	1.42	1.44
4	2.84	2.84	1.87	1.94	2.81	2.85	1.52	1.52
5	2.92	2.91	2.01	2.04	2.83	2.88	1.62	1.61
6	3.05	3.02	2.15	2.16	2.91	2.96	1.72	1.71
7	3.20	3.16	2.29	2.29	3.00	3.05	1.82	1.81
8	3.37	3.31	2.43	2.42	3.12	3.16	1.92	1.91

Table 2.15: In the Table we report the RMSE ratios for the evaluation of the aggregated IRFs, with sum-aggregation scheme. We refer to the Monte Carlo designs: (1) $vec(A) = (0.75, 0.20, 0.10, 0.70)'$, $vech(B) = (1.00, 0.00, 1.00)'$ (on the left), and (2) $vec(A) = (0.75, 0.20, 0.10, 0.70)'$, $vech(B) = (0.60, 0.30, 0.40)'$ (on the right). For each (quarterly) horizon, we compare (i) the aggregated Impulse Response Functions of both the mixed frequency procedures, and (ii) the quarterly IRFs (*naive* approach), relative to the monthly IRFs. Values near to one, indicate that the mixed frequency (or the quarterly frequency) impulse responses are close to the monthly impulse responses. In both the fourth and the eighth column, the value "NaN" is due to the Cholesky restriction.

	$vec(A) = (0.85, 0.20, 0.10, 0.80)'$, $vech(B) = (1.00, 0.00, 1.00)'$				$vec(A) = (0.85, 0.20, 0.10, 0.80)'$, $vech(B = (0.60, 0.30, 0.40)'$			
				State Space vs monthly				
h (quarters)	y to y	x to y	y to x	x to x	y to y	x to y	y to x	x to x
0	1.38	1.56	NaN	1.12	1.47	1.36	NaN	1.24
1	1.12	1.09	1.25	1.07	1.09	1.06	1.31	1.17
2	1.07	1.05	1.20	1.10	1.04	1.04	1.25	1.19
3	1.04	1.04	1.16	1.12	1.02	1.02	1.22	1.19
4	1.02	1.03	1.13	1.11	1.00	1.01	1.20	1.18
5	1.01	1.02	1.10	1.10	1.00	1.00	1.18	1.17
6	1.01	1.01	1.09	1.09	1.00	1.00	1.16	1.16
7	1.01	1.01	1.07	1.08	0.99	0.99	1.14	1.15
8	1.00	1.00	1.06	1.07	0.99	0.99	1.13	1.13
				Minimum Distance vs monthly				
0	1.42	1.70	NaN	1.47	1.52	1.62	NaN	1.44
1	1.11	1.18	1.46	1.05	1.00	1.07	1.12	0.98
2	1.09	1.11	1.40	1.09	0.97	0.99	1.09	0.99
3	1.06	1.07	1.32	1.13	0.98	0.97	1.07	0.99
4	1.04	1.06	1.26	1.14	0.98	0.98	1.06	1.00
5	1.03	1.05	1.21	1.14	0.99	0.98	1.05	1.00
6	1.03	1.04	1.18	1.13	1.00	0.99	1.05	1.02
7	1.04	1.05	1.15	1.12	1.01	1.00	1.05	1.03
8	1.04	1.05	1.14	1.11	1.02	1.01	1.05	1.03
				quarterly vs monthly				
0	8.37	3.95	NaN	7.91	9.04	6.86	NaN	6.17
1	4.33	3.54	1.49	2.91	4.54	4.54	1.11	1.73
2	3.43	3.20	1.49	1.83	3.34	3.45	1.11	1.15
3	2.92	2.88	1.46	1.50	2.72	2.84	1.11	1.06
4	2.59	2.64	1.43	1.40	2.35	2.47	1.12	1.07
5	2.37	2.45	1.40	1.36	2.13	2.23	1.13	1.09
6	2.23	2.31	1.38	1.34	2.00	2.07	1.14	1.11
7	2.14	2.21	1.36	1.33	1.92	1.97	1.14	1.13
8	2.08	2.14	1.35	1.33	1.87	1.91	1.15	1.14

Table 2.16: In the Table we report the RMSE ratios for the evaluation of the aggregated IRFs, with sum-aggregation scheme. We refer to the Monte Carlo designs: (1) $vec(A) = (0.85, 0.20, 0.10, 0.80)'$, $vech(B) = (1.00, 0.00, 1.00)'$ (on the left), and (2) $vec(A) = (0.85, 0.20, 0.10, 0.80)'$, $vech(B) = (0.60, 0.30, 0.40)'$ (on the right). For each (quarterly) horizon, we compare (i) the aggregated Impulse Response Functions of both the mixed frequency procedures, and (ii) the quarterly IRFs (*naive* approach), relative to the monthly IRFs. Values near to one, indicate that the mixed frequency (or the quarterly frequency) impulse responses are close to the monthly impulse responses. In both the fourth and the eighth column, the value "NaN" is due to the Cholesky restriction.

2.5 Generalization to higher lags: sketch of the idea

In the previous sections we have analysed the mapping between high frequency VAR(1) matrices and their low frequency counterparts, obtained either through point-in-time, or through sum-over the low frequency aggregation scheme. By exploiting the analytical relationships found above, we have demonstrated how we can recover the high frequency (unobservable) estimates and then the IRFs using the Classical Minimum Distance estimation method.
In the presence of higher-order VARs the mapping between high and low frequency parameters becomes cumbersome. This can be seen by considering a simple example.
Assume that we know that the high frequency DGP is a monthly structural VAR with two lags, defined by

$$Y_t^* = A_1 Y_{t-1}^* + A_2 Y_{t-2}^* + B\varepsilon_t, \qquad \varepsilon_t \sim WN(0, I) \tag{2.32}$$

with B lower triangular. We assume that the quarterly data are obtained through the point-in-time aggregation scheme. Substituting recursively Y_t^* in Eq. (2.32), we obtain the following expression:

$$\begin{aligned} Y_t^* &= A_1 Y_{t-1}^* + A_2 Y_{t-2}^* + B\varepsilon_t \\ &= A_1^3 Y_{t-3}^* + A_1^2 A_2 Y_{t-4}^* + A_1 A_2 Y_{t-3}^* + A_2 A_1 Y_{t-3}^* + A_2^2 Y_{t-4}^* + \\ &\qquad + A_1^2 B\varepsilon_{t-2} + A_2 B\varepsilon_{t-2} + A_1 B\varepsilon_{t-1} + B\varepsilon_t, \end{aligned}$$

rewritable as

$$\begin{aligned} Y_t^* = (A_1^3 + A_1 A_2 + A_2 A_1) Y_{t-3}^* + (A_1^2 A_2 + A_2^2) Y_{t-4}^* + \\ + (A_1^2 + A_2) B\varepsilon_{t-2} + A_1 B\varepsilon_{t-1} + B\varepsilon_t. \end{aligned} \tag{2.33}$$

It is seen that the presence of the term $(A_1^2 A_2 + A_2^2)Y_{t-4}^*$ makes it impossible the estimation of Eq. (2.33) on quarterly (skip sampled) data. If we consider to iterate the recursive substitution until the low frequency VAR has the same order of the high frequency VAR, i.e. 2, the solution obtained presents the same problem of Eq. (2.33). Specifically, given

$$\begin{aligned} Y_t^* &= (A_1^3 + A_1 A_2 + A_2 A_1) Y_{t-3}^* + (A_1^2 A_2 A_1^2 + A_2^2 A_1^2 + A_2^3 + A_1^2 A_2^2) Y_{t-6}^* + \\ &+ (A_1^2 A_2 A_1 A_2 + A_2^2 A_1 A_2) Y_{t-7}^* + \\ &+ (A_1^2 A_2 A_1 + A_2^2 A_1) B\varepsilon_{t-5} + (A_1^2 A_2 + A_2^2) B\varepsilon_{t-4} + (A_1^2 + A_2) B\varepsilon_{t-2} + A_1 B\varepsilon_{t-1} + B\varepsilon_t, \end{aligned} \tag{2.34}$$

we can note that Eq. (2.34) includes the term $(A_1^2 A_2 A_1 A_2 + A_2^2 A_1 A_2) Y^*_{t-7}$: the instant $t-7$ doesn't coincide to any the quarterly instants $\tau = 3, 6, 9, \dots$.

Assuming that the objects of interest are the IRFs of the high frequency process, a possible solution can be designed as follow. Imagine to know that the high frequency DGP is a VAR(p_m), with $\mathrm{p}_m > 1$ representing the order of the high frequency VAR. However, we can deal only with data sampled at quarterly frequency. Since we can't define a mapping between the high and the low frequency VAR, the first step in the estimation procedure corresponds to the identification of the order of the (observable) low frequency VAR. Using standard information criteria (e.g. Bayesian Information Criterion (BIC), or Akaike Information Criterion (AIC)), we define the order p_q of the low frequency VAR. Then, we estimate the related parameters from the aggregated data and we derive the impulse response functions. In those cases in which p_q is greater then one, the idea is to approximate the IRFs of the estimated VAR(p_q), with IRFs of a quarterly VAR(1). Hence, once obtained the estimates of the VAR(1) which minimize the distance between the related IRFs and the IRFs of the VAR(p_q), the idea is to recover the high frequency impulse responses from the estimates of the quarterly VAR(1). A possible solution to the approximation of the IRFs of the VAR(p_q) can be obtained by referring to the Impulse Response Function Matching estimators[22]. In the literature of DSGE models the idea of the impulse response matching technique is to estimate the structural parameters of a DSGE by minimizing the distance between its impulse responses and the impulse responses obtained with a SVAR(p) model, with $p = 1, 2, \dots$.
In the framework of mixed frequency VARs, focusing on structural analyse, the general idea is to anchor the IRFs obtained from a generic quarterly VAR(p_q) (estimated from the aggregated data and in which the lag p_q is chosen with standard information criteria) to a (misspecified) quarterly VAR(1). By this way, the low frequency estimates obtained from the impulse response matching of the quarterly VAR(1) are merely auxiliary to recover the high frequency IRFs using the proposed approach.
We present below the Impulse Response Function Matching estimation procedure, referring to the classical example provided by the literature, i.e. structural DSGE model estimation. Furthermore, we describe the new approach for the generalization of the proposed procedure and we evaluate the entire estimation technique with a Monte Carlo experiment.

[22] Christiano, Eichenbaum and Evans (2005), Boivin and Giannoni (2006), Fève, Matheron and Sahuc (2009), Hall, Inoue, Nason and Rossi (2012), and Guerron-Quintana, Inoue and Kilian (2017) among others.

2.5.1 Impulse Response Matching Estimation

One of the main aspect in modelling VAR processes is represented by Impulse Response functions. The IRFs and the structural analyses play a strategical role in macroeconomics. Working with a DSGE model, one possible solution to the estimation of its structural parameters has been proposed by Rotemberg and Woodford (1997). Their idea corresponds to obtain the estimates of DSGE structural parameters which lead to IRFs much close as possible to the impulse responses of a SVAR model estimated from the data. The Impulse Response Function Matching Estimator (henceforth, IRFME) proposed by the authors, corresponds practically to a Minimum Distance estimator that minimize the distance between the impulse responses derived from a SVAR(p) model and the impulse responses of a DSGE model.
Consider the generic N-variate VAR(p) process defined by

$$Y_\tau = C_1 Y_{\tau-1} + C_2 Y_{\tau-2} + \cdots + C_p Y_{\tau-p} + R\nu_\tau, \qquad \nu_\tau \sim WN(0, I) \tag{2.35}$$

where Y_τ is a vector of N endogenous variables, R is assumed lower triangular, $R\nu_\tau = \xi_\tau$ and $\Sigma_\xi = RR'$ is the covariance matrix of the residual component ξ_τ. We collect the empirical impulse responses for the horizons $h = 0, 1, \ldots h_{max}$ of the estimated VAR(p) process in the $q \times 1$ vector $\hat{\gamma} = (\hat{\gamma'}_0, \hat{\gamma}'_1, \ldots, \hat{\gamma}'_{h_{max}})'$. The idea is to recover the structural parameter vector ψ of a DSGE model that guarantees the vector of impulse responses $\delta = h(\psi)$ as close as possible to the empirical impulse responses collected in $\hat{\gamma}$. The impulse response function matching estimator is defined by

$$\begin{aligned} \min_{\theta \in \Theta} Q(\theta) &= \min_{\psi \in \Theta_\psi} \{\hat{\gamma} - \delta\}' S \{\hat{\gamma} - \delta\} \\ &= \min_{\psi \in \Theta_\psi} \{\hat{\gamma} - h(\psi\}' S \{\hat{\gamma} - h(\psi)\}, \end{aligned} \tag{2.36}$$

where the weighting matrix S is a definite positive weighting matrix. The optimal choice of S is represented by the inverse of the asymptotic covariace matrix of $\hat{\gamma}$ (see Fevè, Matheron, Sahuc (2009)). Guerron-Quintanta, Inoue and Kilian (2017) consider the inverse of the bootstrap covariance matrix estimator of the impulse responses[23] collected in $\hat{\gamma}$, i.e. $S = \Sigma_{\gamma^*}^{-1}$. In particular, consider the bootstrap

[23] Guerron-Quintanta, Inoue and Kilian (2017) discuss also the use of a non optimal positive definite diagonal weighting matrix, with entries represented by the reciprocals of the diagonal elements of the bootstrap covariance matrix estimator of the impulse responses Σ_{γ^*}. Furthermore, the authors asses that optimal weighting matrix Σ_{γ^*} has a nonstandard convergence rate, and a nonstandard asymptotic distribution when the number of parameters in $\hat{\gamma}$ is higher than the number of the VAR(p) parameters. The solution considered is to mimic the convergence rate and the asymptotic distribution with the bootstrap.

estimator $\hat{\gamma}_b^*$, with $b = 1, \dots, B$, and $\bar{\hat{\gamma}}^* = 1/B \sum_{b=1}^{B} \hat{\gamma}_b^*$. The bootstrap covariance matrix estimator of the impulse responses is given by

$$\hat{\Sigma}_{\gamma^*} = \sum_{b=0}^{B} (\hat{\gamma}_b^* - \bar{\hat{\gamma}}^*)'(\hat{\gamma}_b^* - \bar{\hat{\gamma}}^*).$$

By this way we obtain DSGE structural parameters vector $\hat{\psi}$ that leads to impulse responses (i.e. $\delta = h(\hat{\psi})$) as close as possible to the impulse responses $\hat{\gamma}$ obtained with a VAR(p) model. Standard inference is guaranteed by the order condition $q \leq k$, whit q the number of elements in $\hat{\gamma}$ and k the number of free parameters in the VAR(p) (i.e. $Np + N(N+1)/2$). The optimal choice about the length of $\hat{\gamma}$ is discussed by Hall, Inoue, Nason and Rossi (2012). The authors provide information criteria for the choice of the horizon $\hat{h}_{max}$ of the IRFs in $\hat{\gamma}$. Fevè, Matheron, Sahuc (2009) consider the Redundant Impulse Response Selection Criterion (RIRSC) presented by the working paper of Hall, Inoue, Nason and Rossi (2008). The criterion is a measure of both the numbers k and q, and, in particular, is given by

$$h_{max} = \arg \min_{h_{max} \in \mathcal{H}} = \left\{ log(|\Sigma_\psi|) + \frac{q \, log(T)}{T} \right\},$$

with Σ_ψ the covariance matrix of the estimated parameter vector $\hat{\psi}$. The estimate of $\hat{\Sigma}_\psi$, can be recovered either with bootstrap techniques (see Fevè, Matheron, Sahuc (2009) and Guerron-Quintanta, Inoue and Kilian (2017)), either referring to the asymptotic distribution of Classical Minimum Distance Estimator. This last alternative can be considered only in the case of $q \leq k$.

For the purposes of this discussion, the IRFME provides the quarterly estimates of the matrices C and Ω of Eq. (2.5) from a generic quarterly VAR(p_q) estimated from the data. This result is only auxiliary, and it is used in the mappings identified above for the derivation of the monthly impulse responses. In particular, the idea is to estimate a VAR(p_q) from the aggregated data, with p_q chosen with the BIC information criterion. We derive the impulse responses from the estimates of the VAR(p_q), considering a Cholesky decomposition of the estimated residual covariance matrix. We collect these impulse responses in the vector $\hat{\gamma} = (\hat{\gamma}'_0, \hat{\gamma}'_1, \dots, \hat{\gamma}'_{h_{max}})'$, with $h_{max} = 2$, in accordance with the order condition $q \leq k$ defined by Guerron-Quintana, Inoue and Kilian (2017). The vector of quarterly IRF is given by

$$\hat{\gamma} = \begin{pmatrix} \hat{\gamma}_0 \\ \hat{\gamma}_1 \\ \hat{\gamma}_2 \end{pmatrix},$$

with bootstrap covariance matrix estimator of impulse responses $\hat{\Sigma}_{\gamma^*}$ obtained following Guerron-Quintana, Inoue and Kilian (2017) as described above (with

1000 bootstrap replications). Assume that the vector of parameters of interest is $\psi = (vec(C)', vech(D)')'$, with D the lower triangular matrix obtained from Cholesky decomposition of the residual covariance matrix Ω of the quarterly VAR(1). The mapping between $\hat{\gamma}$ and ψ is given by

$$\begin{aligned}\hat{\gamma}_0 &= vech(D)\\ \hat{\gamma}_1 &= vec(CD)\\ \hat{\gamma}_1 &= vec(C^2D)\end{aligned}$$

Once obtained the estimates of the impulse response matching estimator $\hat{\psi}$, as defined in Eq. (2.36), the proposed Classical Minimum distance procedure of section 2.3.1 can be applied[24]. In this case the optimal weighting matrix S is given by the inverse of the covariance matrix of the estimated vector $\hat{\theta}$, i.e.

$$\hat{\Sigma}_{\psi} = \left(\left(\frac{\partial h(\psi)}{\partial \hat{\psi}'}\right)' \hat{\Sigma}_{\gamma^*}^{-1} \left(\frac{\partial h(\psi)}{\partial \hat{\psi}'}\right)\right)^{-1}.$$

Below we investigate the proposed approach for the generalization to higher order high frequency VARs, with a Monte Carlo experiment.

2.5.2 Monte Carlo experiment

In the description of the Monte Carlo exercise we use the apex $^{(\dagger)}$ to distinguish between the parameters/error terms of the DGP or the parameters/error terms obtained from the data, and the estimates of the auxiliary processes (the quarterly and the monthly VAR(1)).

Assume that the DGP is given by a monthly trivariate ($n = 3$) SVAR(5) process (p_m=5), defined by

$$Y_t^* = A_1^{(\dagger)} Y_{t-1}^* + A_2^{(\dagger)} Y_{t-2}^* + A_3^{(\dagger)} Y_{t-3}^* + A_4^{(\dagger)} Y_{t-4}^* + A_5^{(\dagger)} Y_{t-5}^* B^{(\dagger)} \varepsilon_t^{(\dagger)}, \qquad \varepsilon_t^{(\dagger)} \sim WN(0, I), \tag{2.37}$$

where Y_t^*, $t = 1 \ldots, T$, is the $n \times 1$ vector of the n monthly series, $A_i^{(\dagger)}$, $i = 1, \ldots, 5$, is $n \times n$ coefficient matrix of the i-th lag of the VAR, $u_t^{(\dagger)} = B^{(\dagger)} \varepsilon_t^{(\dagger)}$ is the $n \times 1$ vector of the residuals with covariance matrix $\Sigma^{(\dagger)} = B^{(\dagger)} B^{(\dagger)'}$. $\varepsilon_t^{(\dagger)}$ is the $n \times 1$

[24] The IRFME procedure could be used also for the case presented in section 2.4. With a quarterly VARMA(1,1) solution, under regularity conditions (in particular, stable AR coefficient matrix and invertible MA parameter matrix), we could think to obtain the estimates of the VARMA process referring to a quarterly VAR(∞) with appropriate truncation, VAR(p_q). The following steps coincide with the steps described in this section.

vector of the shocks, and $B^{(\dagger)}$ is the lower triangular coefficient matrix of instantaneous shocks obtained by Cholesky decomposition of $\Sigma^{(\dagger)}$.
For $R = 1000$ replications, we generate the monthly trivariate SVAR(5) process defined in Eq. (2.37) with Y_0 set to $0_{n\times 1}$ and sample size $T = 600$ (months). We assume that at least one of the $n = 3$ variables is sampled at quarterly frequency. By this way, we aggregate the monthly series with point-in-time aggregation scheme. We define the order of the quarterly VAR(p_q) from the aggregated data, i.e.

$$Y_\tau = C_1^{(\dagger)} Y_{\tau-1} + \cdots + C_{p_q}^{(\dagger)} Y_{\tau-1} + \xi_\tau^{(\dagger)}, \qquad \xi_\tau^{(\dagger)} \sim (0, \Omega^{(\dagger)})$$

and we estimate its parameters with least square estimation. $C_j^{(\dagger)}$, with $j = 1, \ldots, p_q$ is the j-th coefficient matrix of the quarterly VAR, $\Omega^{(\dagger)}$ is covariance matrix of the quarterly residuals in $\xi_\tau^{(\dagger)} = D\zeta^{(\dagger)}$, and $D^{(\dagger)}$ is the lower triangular matrix of the coefficients of quarterly instantaneous shocks collected in $\zeta^{(\dagger)}$. Specifically, D is obtained with the Cholesky decomposition of $\Omega^(\dagger)$. The related IRFs are obtained and then the responses for the horizons $h = 0, 1, 2$ are collected in a vector $\hat{\gamma}$. For the estimation of the covariance matrix of the estimated $\hat{\gamma}$, we consider the bootstrap technique[25] described by Guerron-Quintana, Inoue and Kilian (2017). In particular, the bootstrap covariance matrix estimator of the impulse responses is given by

$$\hat{\Sigma}_{\gamma^*} = \sum_{b=0}^{B} (\hat{\gamma}_b^* - \bar{\hat{\gamma}}^*)'(\hat{\gamma}_b^* - \bar{\hat{\gamma}}^*),$$

where $B = 1000$ is the number of bootstrap replications, $\hat{\gamma}_b^*$, with $b = 1, \ldots, B$, is the b-th bootstrap estimator of impulse responses and $\bar{\hat{\gamma}}^* = 1/B \sum_{b=1}^{B} \hat{\gamma}_b^*$. We fix $S = \hat{\Sigma}_{\gamma^*}^{-1}$, and we obtain the quarterly parameter vector $\hat{\theta}$ of a VAR(1) with the impulse response matching technique. Specifically we obtain the estimates of the misspecified quarterly process

$$Y_\tau = CY_{\tau-1} + \xi_\tau, \qquad \xi_\tau \sim (0, \Omega).$$

With the impulse response matching estimator $\hat{C}$ and $\hat{\Omega}$ we obtain the estimates of a quarterly VAR(1) which are as close as possible to the IRF obtained from the data. Even if the application of the Classical Minimum Distance estimation approach described in section 2.3.1, doesn't lead to obtain the estimates of the high frequency VAR(5), the results obtained allows the researcher to derive the estimate of the impulse responses of the DGP. Specifically from the Classical Minimum Distance estimators $\hat{A}$ and $\hat{B}$, defined by

$$Y_t^* = AY_t^* + B\varepsilon_t, \qquad \varepsilon_t \sim (0, \Sigma),$$

[25] A possible alternative solution is proposed by Lutkepohl (1990).

we calculate the high frequency IRFs and we compare the results with the DGP impulse responses.

Results

Since we are not interested in the auxiliary estimates of the quarterly and monthly VAR(1), we evaluate graphically the estimated impulse responses, over 4 years (16 quarters, 48 months). In particular we can consider to divide the analyses of the results in two part:

- the evaluation of the impulse responses obtained with the IRFME;
- the comparison of the high frequency IRFs with the impulse responses obtained with the DGP matrices.

Specifically, in the first evaluation we compare the mean across replications of the IRFs estimated with the quarterly VAR(p_q) with the mean across replication of the IRFs of the VAR(1) obtained after the IRFME. In the proposed exercise, we consider to choose the order of the VAR(p_q) referring to the BIC criterion: over the $R = 1000$ replications of the Monte Carlo simulation, we have estimated (from the data) 559 times a VAR(1), 415 VAR(2) models and 26 VAR(3). In the first case (i.e. the replications in which we select a VAR(1)) we don't need to consider the IRFME: we directly apply the procedure described in section 2.3.1. For the remaining replications we need to involve the impulse response matching approach. In Figure 2.8 we plot the mean across replications of the impulse responses used for the final step of the Classical Minimum Distance procedure. Specifically, we compare the IRFs of the VARs estimated from the data (including the cases in which the information criterion allows us to select the VAR(1)), reported in the figure with the black solid line, and the IRFs obtained from the VAR(1) parameters used for the high frequency IRFs estimation (red solid line). The confidence bounds (red dashed lines) are obtained as two-standar error bounds, with standard errors obtained considering the square root of the diagonal entries of the asymptotic covariance matrix of the IRFs, derived after IRFME.
For the evaluation of these first results we consider also to verify the reliability of the impulse response matching estimates plotting the results obtained for the 441 replications in which the BIC doesn't select the VAR(1) from the aggregated data. In particular, in Figure 2.9 we plot the mean across replications of the impulse responses obtained with a VAR(2) with the related VAR(1) impulse responses derived after IRFME estimation. As in Figure 2.8, we report the IRFs of the (441) VAR(2) and VAR(3) models estimated from the data (black solid line) and the correspondent (441) IRFs obtained with the VAR(1) IRFME estimates (red solid line). The related confidence bounds (red dashed lines) are calculated as

two-standar error bounds, with standard errors obtained as the square root of the diagonal entries of the estimated asymptotic covariance matrix of the estimated IRFs of the VAR(1) solutions. In both the figures, we depict the response of the i-th variable to a shock on the j-th variable as $(j \rightarrow i)$ i.e. (*impulse* $\rightarrow$ *response*). As we can note from both the figures, and in particular in Figure 2.8, the impulse responses obtained from the auxiliary quarterly VAR(1) IRFME estimates seem to be quite accurate and allow us to consider the final minimum distance step for the estimation of the IRFs of the high frequency VAR.

The second kind of the evaluation of the Monte Carlo experiment, consists in comparing the high frequency IRFs (final result of the exercise) with the impulse responses obtained with the DGP matrices.
The first consideration refers to the matrix of the quarterly VAR(1): differently from the previous Monte Carlo experiments (in which only in few replications we had to evaluate the cube root of the quarterly matrix $\hat{C}$ with one real eigenvalue and a pair of complex conjugate eigenvalues) for each replication we estimate the (auxiliary) high frequency parameter matrices $\hat{A}$ and $\hat{B}$ evaluating the three distinct solution of $C^{1/3} \in \mathbb{R}$, and defined by $\{A_1, A_2, A_3\}$ (see section 2.3.1).
In Figure 2.10 we plot the mean across replications of the high frequency impulse responses obtained after the Classical Minimum Distance estimation of $\hat{A}$ and $\hat{B}$ (red solid line) and the IRFs calculated with the DGP matrices of the VAR(5) (black solid line). The confidence bounds (red dashed lines), related to the Minimum Distance estimation, are calculated as two-standar error bounds, with standard errors obtained as the square root of the diagonal entries of the estimated asymptotic covariance matrix of the estimated IRFs of the high frequency VAR(1) solution.
Figure 2.10 shows the ability of the generalized procedure to the estimation of the impulse responses of a high frequency VAR(p$_m$) process, over 48 (monthly) horizons (4 years). Only in the case of the response of the first variable to an impulse in the second, we highlight some differences in the firsts responses with respect to the true values, in particular in the response of the first variable at the horizon $h = 0$.

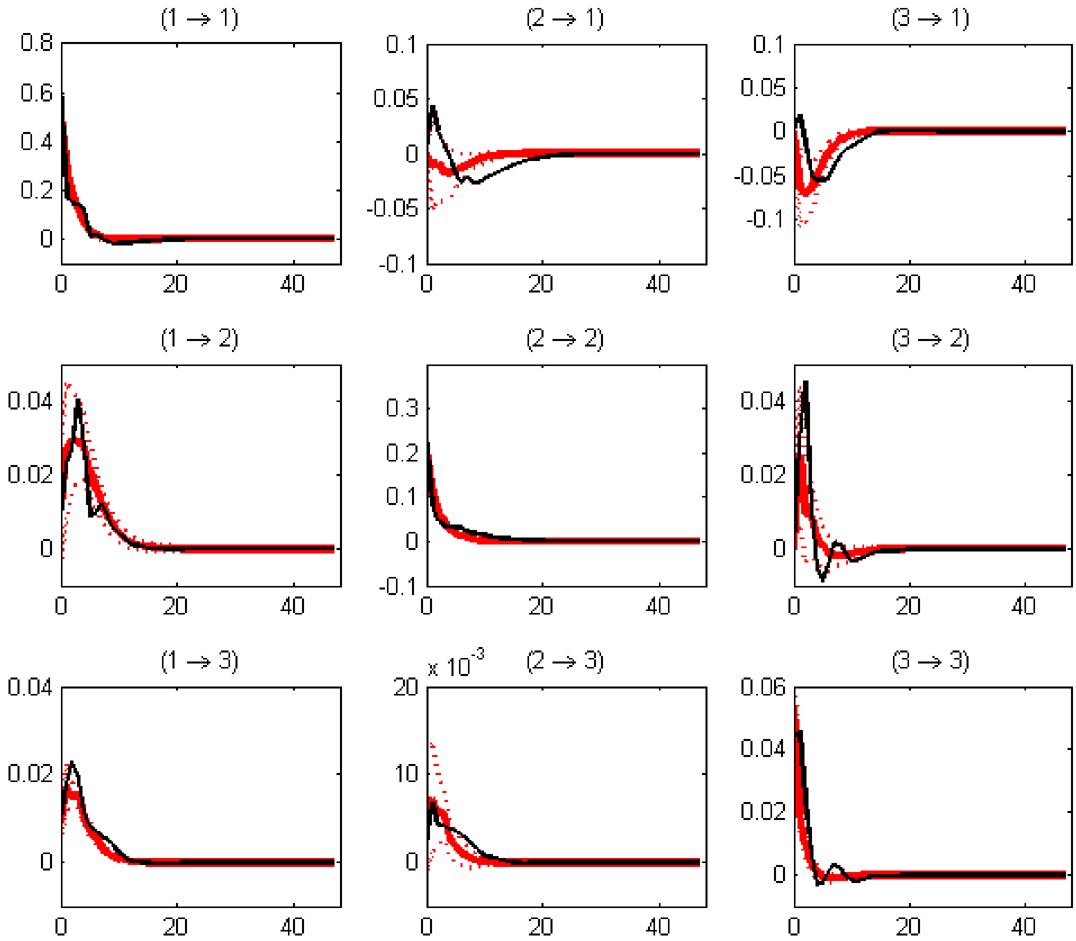


Figure 2.8: In the figure we report IRFs obtained from the data (black solid line) and IRFs obtained after IRFME (red solid line). Both the quantities are obtained as the mean across (R=1000) replications, i.e. the plotted impulse responses are calculated without refer to the different choice of the lag order of the quarterly VAR estimated from the data. The confidence bounds (red dashed lines) are calculated as two-standar error bounds. The standard errors are obtained as the square root of the diagonal entries of the estimated asymptotic covariance matrix of the estimated IRF of the VAR(1) solutions.

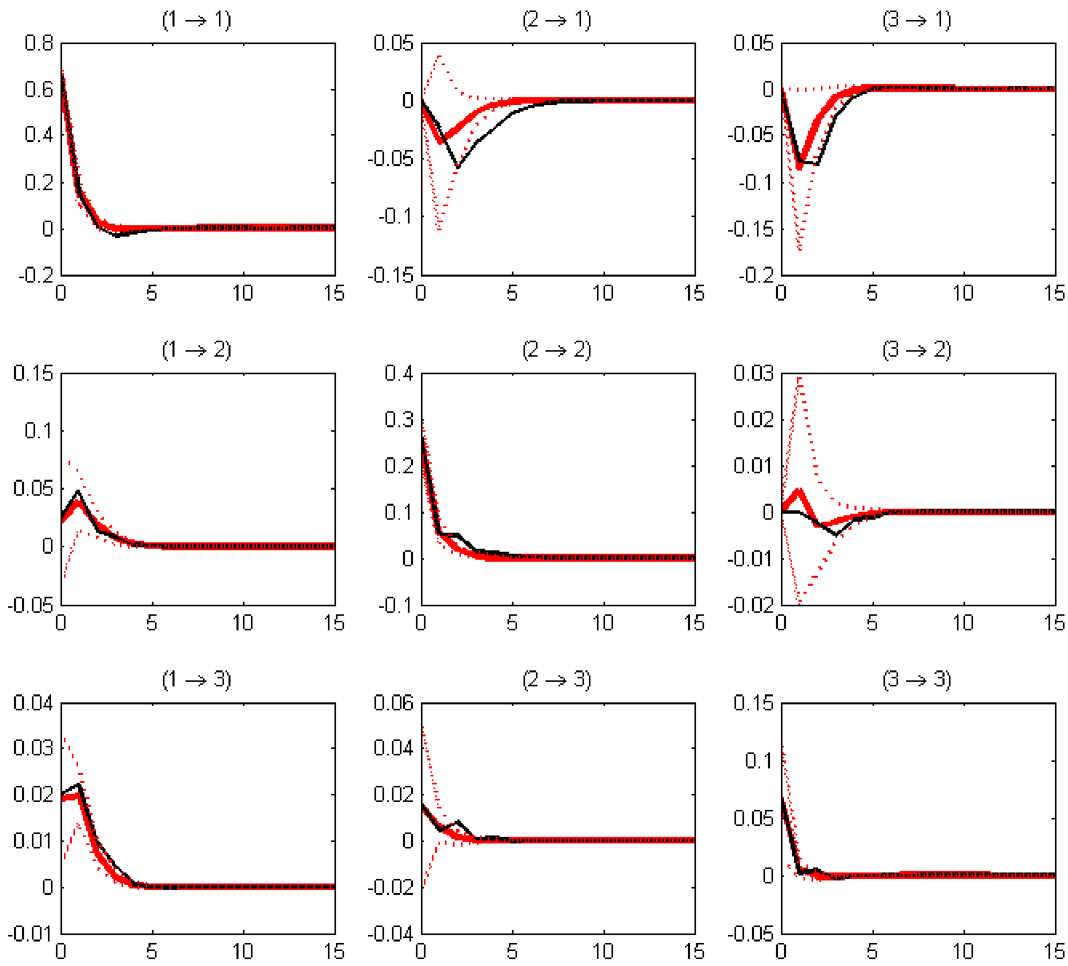


Figure 2.9: In the figure we report IRFs obtained from the data (black solid line) and IRFs obtained after IRFME (red solid line). Both the quantities are obtained as the mean across 441 replications in which the BIC criterion doesn't select the VAR(1) from the aggregated data. The confidence bounds (red dashed lines) are calculated as two-standar error bounds. The standard errors are obtained as the square root of the diagonal entries of the estimated asymptotic covariance matrix of the estimated IRF of the VAR(1) solutions.

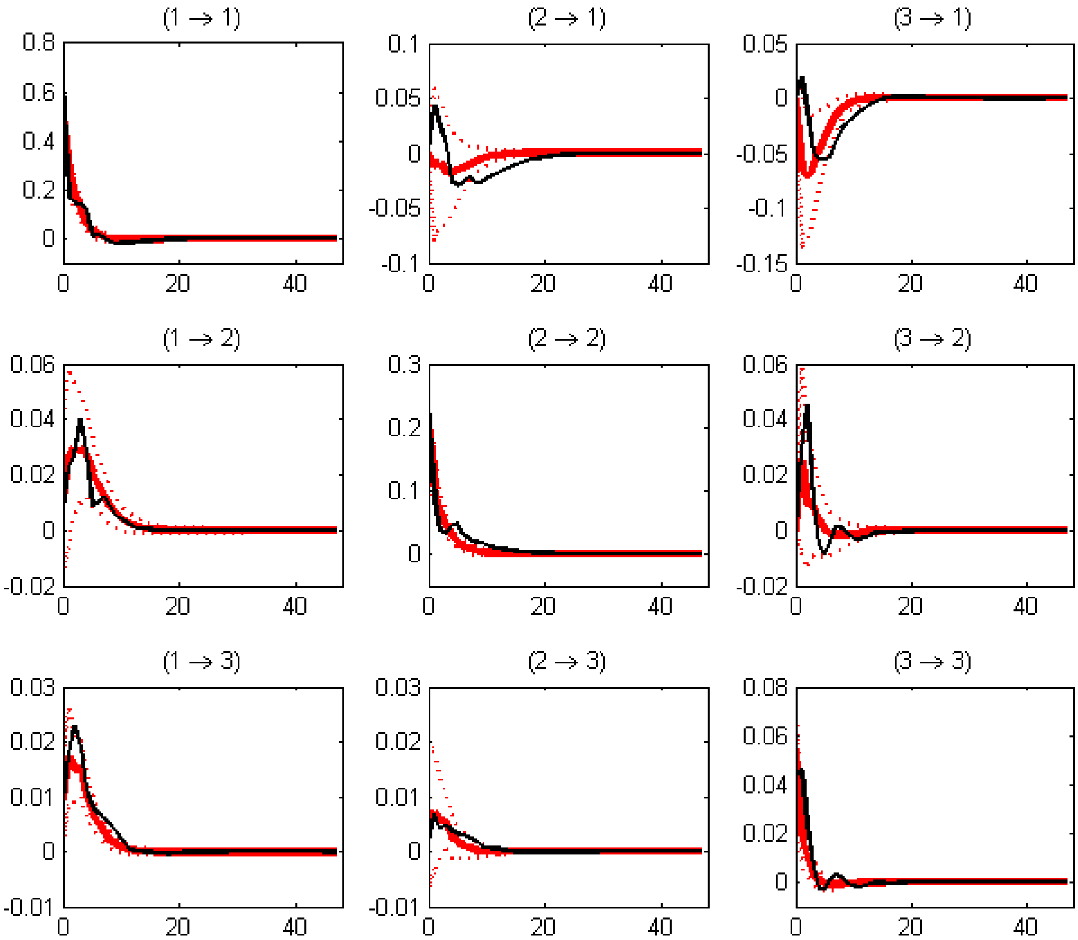


Figure 2.10: In the figure we report IRFs obtained from the DGP matrices of the VAR(5) in Eq. (2.37) (black solid line) and IRFs obtained with the whole procedure (red solid line). The confidence bounds (red dashed lines) are calculated as two-standar error bounds. The standard errors are obtained as the square root of the diagonal entries of the estimated asymptotic covariance matrix of the estimated IRF of the high frequency VAR(1) solutions.

2.6 Empirical Illustration

In empirical macroeconomic applications, referring to datasets composed by variables sampled at different frequency is an usual situation. To evaluate the proposed Minimum Distance estimation procedure we refer to a classical dataset of macroeconomic U.S. variables: the GDP growth rate, as low frequency variable, the Federal Reserve Fund rate (FFR) and the Consumer Price Index growth rate (CPI). FFR and CPI are monthly variables representative of interest and inflation rate respectively, while the GDP growth rate is considered to mimic the output growth. This simple and very classical exercise allows us to evaluate our estimation approach and the standard Kalman filter-based estimation procedure, just considered in the literature.
The comparison of the results consist of two different exercises. In particular, the first part of the empirical application focuses on the comparison of the aggregated IRFs resulting from both the mixed frequency estimation approaches with the quarterly estimates. The objective of this part is to show how the proposed approach could alleviate the aggregation bias.
In the second exercise the mixed frequency results, obtained with each estimation approach, are compared also with a monthly SVAR, in which we consider the Industrial Production growth (IP) as proxy of the output. In this second part of the empirical application we are able to evaluate the the reliability of IRFs obtained with the proposed approach.
The dataset refers to the temporal interval January 1968 - December 2007. As in Foroni and Marcellino (2016) we intentionally exclude the period of the crisis, in order to refer to a stable period. In both the exercises we consider a recursive structural identification scheme.

2.6.1 Results of Exercise 1: mixed frequency vs low frequency

In this empirical exercise we consider the comparison of the solutions obtained with a quarterly dataset and the mixed frequency dataset, from January 1968 to December 2007. In the first step we define and estimate the quarterly VAR(p_q) process from the aggregated data. Considering the naive approach, we aggregate the monthly variables, until the dataset is at the same low frequency. The aggregation scheme chosen is the sum-over-the quarter aggregation scheme. Then, the quarterly vector of endogenous variable is defined by $Y_\tau = (y_\tau, \pi_\tau, r_\tau)'$, where y_τ is the GDP growth rate (as measure of the output growth), π_τ is the CPI growth rate (as inflation rate) and r_τ represents the Federal Fund Rate (as measure of interest rate).
We apply the Cholesky decomposition to make shocks orthogonal, and we trace

out the effects of impulses up to 16 quarters, corresponding to four years. The monthly responses, obtained from the mixed frequency solutions, are computed for 48 periods ahead, to span the the same period of the quarterly process. We depict the response of the i-th variable to a shock on the j-th variable as $(j \rightarrow i)$ i.e. $(impulse \rightarrow response)$.
The monthly results obtained from the mixed frequency approaches are aggregated considering that the shocks are observed in the first month of the quarter, i.e. we aggregate the first, the second and the third monthly responses, in order to compare the high and the low frequency solutions[26].
The aggregated mixed frequency IRFs from the MF-SVAR(5) are reported in Figure 2.11. with the blue and the red solid line. In particular the blue lines represent the aggregated responses obtained after the estimation approach described in section 2.5, and with the red line we show the aggregated responses obtained with the state space approach. The mixed frequency IRFs derived by applying the two approaches are compared with the impulse responses of a quarterly VAR(2). The order of the low frequency VAR is defined referring the Akaike information criterion.
The aggregated solutions and the pure quarterly VAR responses are in line with the literature of temporal aggregation bias (see Marcellino (1999) and Foroni and Marcellino (2016)): in many cases, the magnitude of the effects in the first responses of the quarterly VAR appear greater and more persistent then those obtained with the mixed-frequency approaches.

2.6.2 Results of Exercise 2: mixed frequency vs monthly frequency

In this second exercise, following Foroni and Marcellino (2016), we replace the (quarterly) GDP with the Industrial Production growth rate, available at monthly frequency. Even if the authors, referring only to the analysis of monetary policy shocks, highlight some differences in the responses of the system using either GDP or IP, they conclude that the curves of the responses are sufficiently similar.
In Figure 2.12 we compare monthly IRFs. Specifically, we plot the impulse responses obtained with a monthly SVAR(5) (obtained substituting the GDP growth rate with the Industrial Production growth rate), and the IRFs calculated after the

[26] A different result could be obtained if we imagine that the shocks (the impulses) are observed in the third month of the reference quarter. Specifically, thinking about this situation, even if for the quarterly estimation approach we don't care about the month of the reference quarter in which we observe the shock, for the mixed frequency this aspect becomes quite important: the month in which we observe the impulse provides a different aggregated initial response. This consideration is easily noticeable in three of the responses of our experiment, specifically in the upper-right side of Figure 2.11.

estimation of the MF-SVAR(5), through both the state space approach (red solid line) and the minimum distance procedure (blue solid line). As in section 2.6.1, we apply the Cholesky decomposition to make shocks orthogonal, and we trace out the effects of impulses up four years (48 months).
The Exercise 2 confirms the result obtained in Exercise 1: the responses derived by Minimum Distance estimation are similar to those obtained with the state space procedure, and in line with the monthly results.

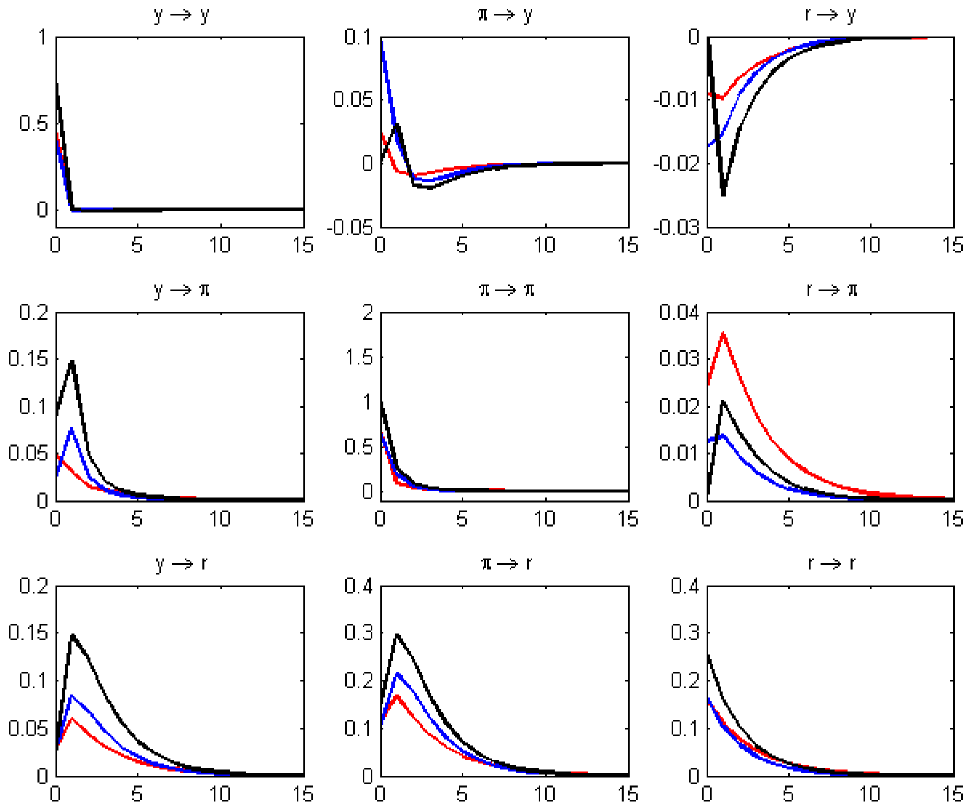


Figure 2.11: Impulse responses obtained from quarterly data (with GDP growth rate as proxy of output growth) in black solid line and the aggregated IRFs from mixed frequency datasets: the low frequency IRFs obtained with the state space approach (red solid line), and aggregated IRFs obtained with Minimum Distance estimation (blue solid line). For each of the three specifications we consider a Cholesky identification scheme. The dataset covers the period 1968q1 - 2007q4.

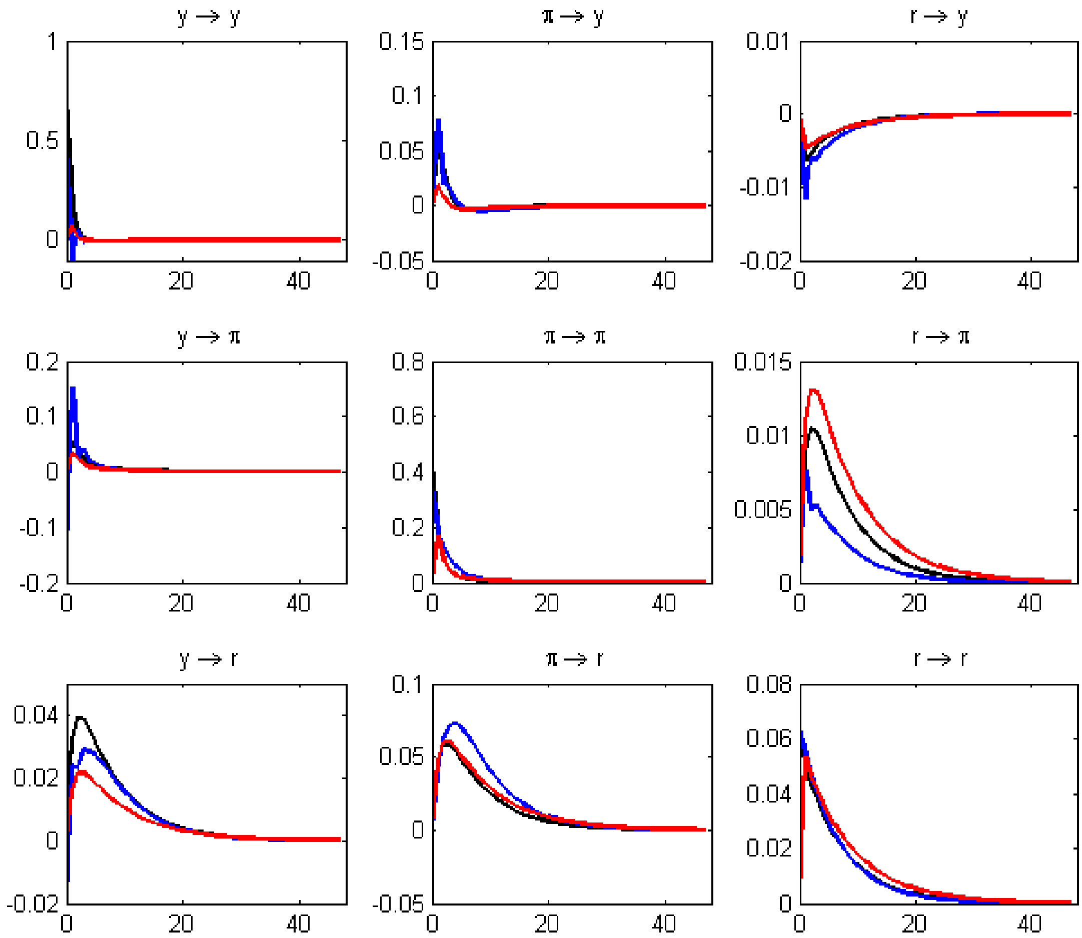


Figure 2.12: Impulse responses at monthly frequency, obtained with the Industrial Production growth rate as proxy of output growth (in black line), IRFs from mixed frequency datasets, obtained with state space approach in red line, and aggregated IRFs obtained with Minimum Distance estimation in blue line. For each of the three specifications we consider a Cholesky identification scheme. The dataset covers the period 1968q1 - 2007q4.

2.7 Conclusion

The recent attention to mixed frequency data, led researcher to provide models able to deal with dataset composed by variables sampled at different frequency. One of this econometric tool is represented by MF-VAR. In this framework the general idea is to use all the information contained in the data, in particular referring to the state space representation. In order to model different economic phenomena, the literature of MF-VAR appears as a growing field of research. One of the aspects less investigated, is represented by the Structural analysis of MF-VAR (MF-SVAR). In particular, this issue is analysed and investigated by Foroni and Marcellino (2014, 2016). The authors asses the usefulness of mixed frequency data to mitigate the problem of aggregation bias and distortions of interpretation in the structural analysis results.
In this discussion we provide a novel estimation procedure that investigates the mapping between the high frequency process (that the research has in mind), and the aggregated low-frequency counterpart. Following the temporal aggregation literature, we consider the most known aggregation schemes: point-in-time sampling and sum sampling, this last as generalization of the average sampling scheme.
Motivated by the recent literature of Linear Rational Expectation models and DSGE models, we start the analyses considering a monthly frequency SVAR(1) process. The schemes identified above lead to different aggregated quarterly results: in the case of point-in-time sampling the low frequency model is a VAR(1), while with the sum sampling, the aggregated result is a VARMA(1,1). From these solutions, we identify the mapping between the monthly and the quarterly parameter matrices and we estimate the high frequency parameters with Minimum distance estimation.
In the discussion, we provide different Monte Carlo experiments to evaluate the capability of the proposed approach and reliability of the results. For both the aggregation schemes, we compare the minimum distance results with the estimates obtained with the state space procedure. Moreover, if the attention of the researcher is focused merely on the estimation of the impulse response functions, we propose a possible generalization of the procedure to higher order (high frequency) VARs referring to the Impulse Response Functions Matching Estimation. The limit of this approach is represented by the impossibility to recover the estimates of the high frequency VAR(p) that we have in mind.
In general, even if, the state space procedure seems to work a little bit better then minimum distance, especially in term of variability of the estimates, the results of the two approaches can be considered comparable.

APPENDIX

A1. The Cube Root of a Matrix

In this Appendix we consider the problem of computing the cube root of a square matrix C, i.e. $C = A^3$, where C is known and the problem is computing A.
In general, a function $f : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ of a matrix C can be searched by different procedures. Higham (2008) summarizes the most useful ones: the Jordan canonical form, the polynomial interpolation and the Cauchy integral. We focus on the method based on the Jordan Canonical form, since it will be used in the paper. Before addressing the specific problem of finding a cube root of a matrix, we introduce some definitions.

Definition 1. Matrix functions defined by Jordan canonical form: *Let $C \in \mathbb{C}^{n\times n}$ an $n \times n$ matrix. Its Jordan canonical form is given by*

$$M^{-1}CM = J = \begin{bmatrix} J_{k_1}(\lambda_1) & 0 & \cdots & 0 \\ 0 & J_{k_2}(\lambda_2) & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & J_{k_r}(\lambda_r) \end{bmatrix}$$

where M is an $n \times n$ nonsingular matrix, λ_i, with $i = 1, \dots, r$, $r \leq n$, are the eigenvalues of C, $J_i(\lambda_i) \equiv J_i \in \mathbb{C}^{k_i\times k_i}$ is the ith Jordan block, given by:

$$J_{k_i}(\lambda_i) \equiv J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & \lambda_i \end{bmatrix} \in \mathbb{C}^{k_i\times k_i},$$

where k_i is the number of repeated eigenvalue λ_i and $k_1 + k_2 + \cdots + k_r = n$.
Then the matrix function $f(C)$, with $f : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$, is defined by

$$f(C) = f(MJM^{-1}) = Mf(J)M^{-1} = Mdiag(f(J_1), \dots, f(J_r))M^{-1}. \quad \text{(A1.1)}$$

In Eq. (A1.1)*, given* $f_0 \equiv f : \mathbb{C} \to \mathbb{C}$*,* $f(J_i)$ *is given by:*

$$f(J_i) = \begin{bmatrix} f_0(\lambda_i) & f_0^{(1)}(\lambda_i) & \cdots & \frac{f_0^{(k_i-1)}(\lambda_i)}{(k_i-1)!} \\ 0 & f_0(\lambda_i) & \ddots & \vdots \\ \vdots & \ddots & \ddots & f_0^{(1)}(\lambda_i) \\ 0 & \cdots & 0 & f_0(\lambda_i) \end{bmatrix},$$

where $f_0^{(d)}$ *represents the dth derivative of* f_0*.*

The problem we focus on, is only a particular case of a wide branch of the mathematical literature. Considering f multivalued[27], we have to evaluate several aspects of the problem: first of all, the *derogatory* - *nonderogatory* nature of the matrix C. When the Jordan matrix presents repeated eigenvalues (i.e. C is a derogatory matrix), we can consider the same branch[28] for f and its derivatives, or different branches for each repeated eigenvalues. In the first case we evaluate the *primary matrix function*, and in the second case we consider the *non primary* solution.
The function f we are interested in is the cube root of a matrix $C \in \mathbb{R}^{3\times 3}$, with distinct eigenvalues (i.e. C is a *non derogatory* matrix)[29]. In this case, f is a multivalued function and consider $f : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ with $n = 3$.
Depending on the eigenvalues of C, we identify two cases, denoted Case 1 and Case 2, respectively:

- Case 1. C has three distinct eigenvalues in $\mathbb{R}$;

- Case 2. C has one eigenvalue in $\mathbb{R}$ and a pair of complex conjugate eigenvalues (in $\mathbb{C}$).

From Definition 1 it follows that our problem amounts to finding the cube roots of each eigenvalue.

[27]For one point in its domain, a multivalued functions assumes at least two distinct values in its range. In this specific case, the cube root of a number is a multivalued function: for one point in its domain, the cube root (f_0) assumes three distinct values in its range.

[28]The term *branch* is related to multivalued functions. In the complex domain, a branch (also called, *sheets*) is a portion of the range of a multivalued function over which the function is single-valued. In particular, combining all the branches gives the full structure of the function. A *principal branch* is obtained when a particular branch of a function is chosen to work with.

[29]In this case, the Jordan blocks are scalars, implying all the elements of the superdiagonals of J are equal to zero.

Case 1: Unique Real Solution of the Cube Root of C

In Case 1, we observe three distinct eigenvalues in $\mathbb{R}$. In this case exists a unique solution $A \in \mathbb{R}^{3\times 3}$, which solves $C^{1/3} = A$. Consider the eigenvalues of C, $\{\lambda_1, \lambda_2, \lambda_3\} \in \mathbb{R}$. For each λ_i, $i = 1, 2, 3$, the number of cube roots is three: one in the real plane, defined as r, and two in the complex plane, specifically, a conjugate pair $\{c, \bar{c}\}$, where $\bar{c}$ is the complex conjugate of c.
For instance, the cube roots of $\lambda_i = -8$ are:

$$\sqrt[3]{-8} = \begin{cases} -2.0000 \\ 1.0000 + 1.7321i \\ 1.0000 - 1.7321i \end{cases},$$

where $i = \sqrt{-1}$ is the imaginary unit.
Then, from the 3×3 Jordan matrix J, with eigenvalues $(\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}$, we can calculate

$$\lambda_1^{1/3} = \{r_1, c_1, \bar{c}_1\}, \tag{A1.2a}$$
$$\lambda_2^{1/3} = \{r_2, c_2, \bar{c}_2\}, \tag{A1.2b}$$
$$\lambda_3^{1/3} = \{r_3, c_3, \bar{c}_3\}. \tag{A1.2c}$$

The unique solution $A = C^{1/3} \in \mathbb{R}^{3\times 3}$ is given by:

$$A = C^{1/3} = MJ^{1/3}M^{-1} = M \begin{bmatrix} r_1 & 0 & 0 \\ 0 & r_2 & 0 \\ 0 & 0 & r_3 \end{bmatrix} M^{-1}.$$

Any other combination of the roots of the eigenvalues in Eq. (A1.2), produces cube roots of C, $MJ^{1/3}M^{-1}$ outside the real plane, i.e. with complex entries.

Case 2: Indeterminacy of the Real Solution of the Cube Root of C

In Case 2, does not exist a unique A which solves $C^{1/3} = A \in \mathbb{R}^{3\times 3}$. Consider the eigenvalues of C, $\{\lambda_1, \lambda_2, \lambda_3\}$, where $\lambda_1 \in \mathbb{R}$ and $\{\lambda_2, \lambda_3\} \in \mathbb{C}$ are conjugates, $\lambda_3 \equiv \bar{\lambda}_2$. From the diagonal of the Jordan matrix, we observe:

$$\lambda_1^{1/3} = \{r_1, c_1, \bar{c}_1\}, \tag{A1.3a}$$
$$\lambda_2^{1/3} = \{c_{12}, c_{22}, c_{32}\}, \tag{A1.3b}$$
$$\lambda_3^{1/3} \equiv \bar{\lambda}_2^{1/3} = \{\bar{c}_{12}, \bar{c}_{22}, \bar{c}_{32}\}. \tag{A1.3c}$$

The only (distinct) cube roots of $C \in \mathbb{R}^{3\times 3}$ are obtained through the Jordan matrices

$$J_1^{1/3} = \begin{bmatrix} r_1 & 0 & 0 \\ 0 & c_{12} & 0 \\ 0 & 0 & \overline{c}_{12} \end{bmatrix}, J_2^{1/3} = \begin{bmatrix} r_1 & 0 & 0 \\ 0 & c_{22} & 0 \\ 0 & 0 & \overline{c}_{22} \end{bmatrix}, J_3^{1/3} = \begin{bmatrix} r_1 & 0 & 0 \\ 0 & c_{32} & 0 \\ 0 & 0 & \overline{c}_{32} \end{bmatrix}.$$

Any other combination of the roots of the eigenvalues in Eq. (A1.3), produces cube roots of C, $MJ^{1/3}M^{-1}$, outside the real plane, i.e. with complex entries.

A2. Estimate of the unique monthly matrix $\hat{A}$

In this Appendix, we consider the estimation of A, given the relationship $C = A^3$. Assume that $vec(\hat{C})$ is any consistent, asymptotically Gaussian estimate of $vec(C)$, with covariance matrix $\hat{\Omega}$. In order to estimate $\hat{A} = \hat{C}^{1/3}$, we consider the two cases discussed in APPENDIX A1.
Case 1 does not pose any problem in searching the (unique) solution $\hat{A} = \hat{C}^{1/3} \in \mathbb{R}^{3\times 3}$. Case 2, instead, requires finding a method to select the $\hat{A}_i$, $i = 1, 2, 3$, from the three real equivalent solutions of $\hat{C}^{1/3} = \{\hat{A}_1, \hat{A}_2, \hat{A}_3\}$.

Case 1

- $\hat{C}$ is decomposed in its Jordan canonical form, obtaining the matrix $\hat{M}$ and $\hat{J}$ from the decomposition:

$$\hat{C} = \hat{M}\hat{J}\hat{M}^{-1};$$

- the real cube roots of the (real) entries of $\hat{J}$ is calculated, obtaining $\hat{J}^{1/3}$;
- the resultant diagonal matrix $\hat{J}^{1/3}$ is then pre-multiplied by $\hat{M}$ and post-multiplied by $\hat{M}^{-1}$, obtaining $\hat{A} = \hat{M}\hat{J}^{1/3}\hat{M}^{-1} = \hat{C}^{1/3} \in \mathbb{R}^{3\times 3}$.

Case 2

- $\hat{C}$ is decomposed in its Jordan canonical form, obtaining the matrix $\hat{M}$ and $\hat{J}$ from the decomposition:

$$\hat{C} = \hat{M}\hat{J}\hat{M}^{-1};$$

- we calculate the cube roots of the three Jordan matrices, as illustrated in the APPENDIX A1 - Case 2, $\hat{J}_i^{1/3}$, with $i = 1, 2, 3$; in particular, we consider the three combinations of

  - the real cube root (from Eq. (A1.3), r_1) of the only real entry of $\hat{J}$, with
  - the three pair of complex cube roots of the conjugate pairs (specifically, from Eq. (A1.3), $\{c_{12}, \bar{c}_{12}\}$, $\{c_{22}, \bar{c}_{22}\}$, $\{c_{32}, \bar{c}_{32}\}$);

- each resultant diagonal matrix, $\hat{J}_i^{1/3}$, with $i = 1, 2, 3$, is pre-multiplied by $\hat{M}$ and post-multiplied by $\hat{M}^{-1}$, obtaining three distinct real matrices: $\hat{A}_1, \hat{A}_2, \hat{A}_3$;

In order to select one among the three matrices, we present two selection methods. These procedures guarantee that the selected $\hat{A}$ minimizes the distance between the observations, according to some criteria.

Case 2 - Procedure 1

- For each $\hat{A}_i$, with $i = 1, 2, 3$, calculate the minimum distance estimates of B and V, i.e. $\hat{B}_i$ and $\hat{V}_i$, with $i = 1, 2, 3$;

- from the estimated $\{\hat{A}_1, \hat{B}_1, \hat{V}_1\}, \{\hat{A}_2, \hat{B}_2, \hat{V}_2\}$ and $\{\hat{A}_3, \hat{B}_3, \hat{V}_3\}$ generate the fitted monthly series $Y_t^{\star\star} = (y_{1,t}^{\star\star}, y_{2,t}^{\star\star}, y_{3,t}^{\star\star})'$, with observations obtained as:

$$Y_{i,t}^{\star\star} = \begin{cases} \hat{A}_i Y_{i,t-1}^{*}, & \varepsilon_{i,t} \sim (\mathbf{0}_3, I_3) \quad \text{if } t \neq \tau, t = 1, \ldots T_b \\ \hat{A}_i Y_{i,t-1}^{*}, & \varepsilon_{i,t} \sim (\mathbf{0}_3, \hat{V}_i) \quad \text{if } t \neq \tau, t = T_b + 1, \ldots T \end{cases}$$

  with $i = 1, 2, 3$ and where $\varepsilon_{i,t} \sim (\mathbf{0}_3, I_3)$ and $\varepsilon_{i,t} \sim (\mathbf{0}_3, V_i)$ are obtained from $S = 100$ draws from a multivariate Normal distribution with mean $\mathbf{0}_3$ and covariance matrices: (i) I_3 until T_b, and (ii) V_i for $t = T_b + 1, \ldots, T$;

- we calculate the adjusted R^2 for the monthly fitted variables generated by each set of matrices;

- the chosen matrix $\hat{A}^*$, (hence, $\{\hat{A}^*, \hat{B}^*, \hat{V}^*\}$), will be the one which provides the maximum value of the index.

Case 2 - Procedure 2

- For each estimation we observe the value of the minimization function of the minimum distance estimators;

- the chosen matrix $\hat{A}^*$, (hence, $\{\hat{A}^*, \hat{B}^*, \hat{V}^*\}$), will be the one which provides the minimum value of the function.

A3. Classical Minimum Distance Estimator

In general, the extremum estimation method could be described as a class of estimation methods. The Maximum Likelihood, least squares, Method of Moments and Classical minimum distance estimation are only a few of methods belong to this general class. In particular, in this work, we concentrate our attention on the Classical minimum distance estimation method[30].

Suppose that θ_0 is a $k \times 1$ vector of parameters of interest, which is known to be function of the $h \times 1$ parameter vector ϕ_0, with $h > k$. In particular, for a known continuously differentiable function h,

$$h(\theta_0, \phi_0) = 0. \tag{A3.1}$$

Let $\hat{\phi}$ be a consistent and asymptotically Normal estimator of ϕ_0, with asymptotic covariance matrix Ψ_∞, specifically

$$\sqrt{T}(\hat{\phi} - \phi_0) \sim \mathcal{N}(\mathbf{0}, \Psi_\infty).$$

The minimum distance estimator $\hat{\theta}$ solves the minimization problem:

$$\min_{\theta \in \Theta} Q(\theta) = \min_{\theta \in \Theta} \{h(\theta_0, \hat{\phi}\}' S \{h(\theta_0, \hat{\phi})\}, \tag{A3.2}$$

where S is any positive semi-definite symmetric matrix.
Malinvaud (1970) demonstrates that, since there are no restrictions on the covariance matrix of $\hat{\phi}$, the minimum distance estimator $\hat{\theta}$ is asymptotically efficient in the class of minimum distance estimators when S is equal to the inverse of the asymptotic variance of $h(\hat{\theta}, \hat{\phi})$.
In particular, the covariance matrix can be calculated applying the Delta method[31].

[30] Malinvaud (1970) introduced this class of method, first, as non-linear estimation procedure, starting from Gauss-Markov theory of linear estimation and trying to see what seems different in a non linear structure.

[31] In general, consider any consistent estimator G of the true parameter Γ. By the central limit theorem: $\sqrt{n}(G - \Gamma) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \Sigma)$. Suppose we are interested in estimate the covariance matrix of a function h of G. Considering Taylor series expansion of the second order:

$$h(G) \approx h(\Gamma) + H(\Gamma)'(G - \Gamma)$$

where $H(\Gamma)$ is the gradient of $h(\Gamma)$. Then, the covariance of $h(G)$ is, approximately, given by:

$$\begin{aligned} Var(h(G)) &\approx Var(h(\Gamma) + H(\Gamma)'(G - \Gamma)) = \\ &= Var(h(\Gamma) + H(\Gamma)'G - H(\Gamma)'\Gamma) = \\ &= Var(H(\Gamma)'G) = \\ &= H(\Gamma)' Var(G) H(\Gamma), \end{aligned}$$

implying $\sqrt{n}(h(G) - h(\Gamma)) \xrightarrow{d} \mathcal{N}(\mathbf{0}, H(\Gamma)'\Sigma H(\Gamma))$.

Let

$$H_\phi = \frac{\partial h(\theta,\phi)}{\partial \phi'}, \qquad H_\theta = \frac{\partial h(\theta,\phi)}{\partial \theta'},$$

the Jacobian matrices[32] of the function $h(\theta,\phi)$. Assume that $\hat{\Psi}$ is any consistent estimator of Ψ_∞. By the application of the Delta method, the asymptotic covariance matrix of $h(\theta,\hat{\phi})$, is given by

$$W_{\hat{\phi},\theta} = (H_\theta' S H_\theta)^{-1}\left[H_\theta' S(H_\phi'\hat{\Psi}H_\phi)S'H_\theta\right](H_\theta' S H_\theta)^{-1}.$$

If $S = (H_\phi'\hat{\Psi}H_\phi)^{-1}$, the minimum distance estimator $\hat{\theta}$ is asymptotically efficient in the class of minimum distance estimators, and $W_{\hat{\phi},\hat{\theta}}$

$$W_{\hat{\phi},\hat{\theta}} = (H_\theta'(H_\phi'\hat{\Psi}H_\phi)^{-1}H_\theta)^{-1}.$$

The minimization function $Q(\theta)$ can be defined as:

$$Q(\theta) = h(\theta,\hat{\phi})'\hat{S}_{\phi,\theta}h(\theta,\hat{\phi}),$$

where $\hat{S}_{\phi,\theta} = (H_\phi'\hat{\Psi}H_\phi)^{-1}$.
As we can note from the Jacobian, the weighting matrix $S = \hat{S}_{\phi,\theta}$, is function of θ. A classical approach in minimum distance estimation is a two-step procedure. In the first step we obtain θ^* as the minimum distance estimator of θ_0 with $S = I$. Then $\hat{S}_{\hat{\phi},\theta}$ is obtained as a function of θ^*, in particular, from the Jacobian matrix

$$H_\phi^* = \frac{\partial h(\theta^*,\hat{\phi})}{\partial \hat{\phi}'}.$$

Then, the criterion function $Q(\theta)$ is given by:

$$Q(\theta) = h(\theta,\hat{\phi})'\hat{S}_{\hat{\phi},\theta^*}h(\theta,\hat{\phi}) = h(\theta,\hat{\phi})'(H_\phi^{*\prime}\hat{\Psi}H_\phi^*)^{-1}h(\theta,\hat{\phi}).$$

The procedure described above corresponds to the two-step classical minimum distance estimation.
A particular case of the minimization problem described above is obtained considering the explicit formulation of the restrictions in Eq. (A3.1). For a known continuously differentiable function $f : \mathbb{R}^k \to \mathbb{R}^h$, we observe:

$$\phi_0 = f(\theta_0).$$

[32] As pointed out in Newey and McFadden (1994), standard inference is guaranteed by the full rank of the Jacobian matrices.

In this case, the minimum distance estimator $\hat{\theta}$ is given by:

$$\min_{\theta \in \Theta} Q(\theta) = \min_{\theta \in \Theta} \{\hat{\phi} - f(\theta)\}' S \{\hat{\phi} - f(\theta)\}.$$

with optimal weighting matrix S equals the inverse of the estimator of the asymptotic covariance matrix Ψ_0, i.e. $S = \hat{\Psi}^{-1}$.

In the literature, many different choices of S have been proposed. Brown (1960) and Nakamura (1960), respectively, proposed and demonstrated the convergence of the minimum distance estimator with $S = I$, where I is the identity matrix. In recent years, Rotemberg and Woodford (1997) and Amato and Laubach (2003), use $S = I$ in minimum distance estimation of structural VAR models. Boivin and Giannoni (2006) and Christiano, Eichenbaum and Evans (2005) estimated the VAR structural parameters considering a diagonal matrix with entries equal the inverse of the Impulse Response Function's variances (i.e. CEE-type weighting matrix). Several are also the minimum distance estimation methods implying two step procedures. In addition to the two-step classical minimum distance described above (see Newey and McFadden (1994), for further details), we can refer to S as a function of θ and solve the minimization function *w.r.t.* the parameter vector θ. The estimated $\hat{S}$ is, then, used in solving the minimization problem summarized in Eq. (A3.2). If S and θ are estimated simultaneously, we refer to the Continuously Updating (CU) GMM estimator of Hansen, Heaton and Yaron (1996).

A4. VARMA identification through mixed frequency data

Consider the results obtained in Section 2.2. The aggregation of a monthly SVAR(1) process, defined as

$$Y_t^* = AY_{t-1}^* + B\varepsilon_t, \qquad \varepsilon_t \sim (\mathbf{0}, I_3).$$

provides the quarterly VARMA(1,1) process:

$$\begin{aligned} Y_t^* + Y_{t-1}^* + Y_{t-2}^* =& A^3(Y_{t-3}^* + Y_{t-4}^* + Y_{t-5}^*) + \\ & + B\varepsilon_t + (I + A)B\varepsilon_{t-1} + (I + A + A^2)B\varepsilon_{t-2} + \\ & + (A^2 + A)B\varepsilon_{t-3} + A^2 B\varepsilon_{t-4} \end{aligned}$$

$$Y_\tau = CY_{\tau-1} + \xi_\tau + Q\xi_{\tau-1}, \qquad \xi_\tau \sim (0, \Pi) \tag{A4.1}$$

with

$$\begin{aligned} \xi_\tau &= B\varepsilon_t + (I + A)B\varepsilon_{t-1} + (I + A + A^2)B\varepsilon_{t-2}, \\ Q\xi_{\tau-1} &= (A + A^2)B\varepsilon_{t-3} + A^2 B\varepsilon_{t-4} \end{aligned}$$

and

$$\Pi = BB' + (I + A)BB'(I + A)' + (I + A + A^2)BB'(I + A + A^2)'.$$

Assume $\eta_\tau = \xi_\tau + Q\xi_{\tau-1}$ and

$$\mathbb{E}(\eta_\tau \eta_{\tau-h}) = \begin{cases} \Omega = \Pi + Q\Pi Q' & \text{if } h = 0, \\ \Phi = Q\Pi & \text{if } h = 1, \\ 0 & \text{if } h \geq 2. \end{cases} \tag{A4.2}$$

With the aim of deriving the high frequency parameters, we have to estimate the quarterly counterpart. However, as pointed out in the literature, the estimation of a VARMA(p,q) process presents identification problems. In particular, the uniqueness of the VARMA representation is guarantees by the introduction of some restrictions. Classical choice of identification schemes for VARMA processes are represented by the echelon form and the final equation form: in both the cases we reduce the number of estimated parameters, achieving identification. For further details see Lutkepohl (2012).
For the problem introduced in Eq. (A4.1), we propose a procedure in order to identify and recover the VARMA coefficient matrix from the estimated high frequency parameters $\hat{A}$ and $\hat{B}$.
First, we consider a consistent estimator of the coefficient matrix C, for instance, the Instrumental Variables (IV) estimator $\hat{C}_{IV}$. Then, the estimator of the monthly matrix $\hat{A}$ is calculated as the cube root of $\hat{C}_{IV}$, i.e.

$$\hat{A} = (\hat{C}_{IV})^{\frac{1}{3}}.$$

Given the quarterly residual component $\hat{\eta}_\tau = Y_\tau - \hat{C}_{IV} Y_{\tau-1}$, we can obtain an estimate of the covariance matrix $\hat{\Omega}$ and of the first order autocovariance matrix $\hat{\Phi}(1) \equiv \hat{\Phi}$, respectively as

$$\hat{\Omega} = \frac{1}{T}\hat{\eta}_\tau \hat{\eta}_\tau', \qquad \hat{\Phi} = \frac{1}{T}\hat{\eta}_\tau \hat{\eta}_{\tau-1}'.$$

Given $\hat{\Omega}$ and $\hat{\Phi}$ we can obtain the monthly structural coefficient matrix $\hat{B}$, by Minimum Distance estimation, with recursive and non-recursive structure. Hence, once obtained $\hat{B}$, we can also estimate indirectly the remaining component of the quarterly process, i.e. $\hat{\Pi}$, as a function of the monthly estimates $\hat{A}$ and $\hat{B}$,

$$\hat{\Pi} = \hat{B}\hat{B}' + (I + \hat{A})\hat{B}\hat{B}'(I + \hat{A})' + (I + \hat{A} + \hat{A}^2)\hat{B}\hat{B}'(I + \hat{A} + \hat{A}^2)'.$$

Furthermore, from the relation in Eq. (A4.2) we are also able to recover the quarterly parameter of the coefficient matrix Q. In particular

$$\begin{aligned} \hat{Q} &= \hat{\Phi}\hat{\Pi}^{-1} = \\ &= [(A + A^2)BB' + A^2BB'(I + A)'] \\ &\qquad [\hat{B}\hat{B}' + (I + \hat{A})\hat{B}\hat{B}'(I + \hat{A})' + (I + \hat{A} + \hat{A}^2)\hat{B}\hat{B}'(I + \hat{A} + \hat{A}^2)']^{-1}. \end{aligned}$$

Bibliography

[1] Aastveit, K. A., Gerdrup, K. R., Jore, A. S., Thorsrud, L. A., (2014), *Nowcasting GDP in Real Time: A Density Combination Approach*. Journal of Business & Economic Statistics, 32:1, 48-68.

[2] Amato, J. D., Laubach, T., (2003), *Estimation and Control of an Optimization-Based Model with Sticky Prices and Wages*. Journal of Economic Dynamics and Control 27(7), 1181-1215.

[3] Amemiya, T. ,(1974), *The Non-linear Two-stage Least-Squares Estimator*. Joumal of Econometrics, 2, pp. 105-110.

[4] Amemiya, T., (1977), *The Maximum Likelihood and Nonlinear Three-Stage Least Squares Estimator in the General Nonlinear Simultaneous Equations Model*. Econometrica, 45, pp. 955-968.

[5] Amisano, G., Giannini, C., (1997), *Topics in Structural VAR Econometrics 2nd ed.* Springer, Berlin.

[6] Andrews, D. W. K. (1991), *Heteroskedasticity and autocorrelation consistent covariance matrix estimation*. Econometrica, vol. 59(3), pp. 817-858.

[7] Bacchiocchi, E., Castelnuovo, E., Fanelli, L., (2016), *Gimme a Break! Identification and Estimation of the Macroeconomic Effects of Monetary Policy Shocks in the U.S.*. Macroeconomic Dynamics, forthcoming

[8] Bacchiocchi, E., Fanelli, L., (2015), *Identification in Structural Vector Autoregressive Models with Structural Changes, with an Application to US Monetary Policy*. Oxford Bulletin of Economics and Statistics, 0305-9049, doi: 10.1111/obes.12092.

[9] Baffigi, A., Golinelli, R., Parigi, G., (2004), *Bridge models to forecast the euro area GDP*. International Journal of Forecasting, 20(3), 447-460.

[10] Banbura, M., Giannone, D., Modugno, M. and Reichlin, L. (2013), *Nowcasting and the real time data flow*. Handbook of Economic Forecasting, vol. 2 (eds G. Elliott and A. Timmermann). Amsterdam: North-Holland.

[11] Banbura, M., Modugno, M., (2014), *Maximum Likelihood Estimation of Factor Models on Datasets with Arbitrary Pattern of Missing Data*. Journal of Applied Econometrics, 29, 133-160.

[12] Banbura, M., Rüstler, G., (2011), *A Look into the Factor Model Black Box: Publication Lags and the Role of Hard and Soft Data in Forecasting GDP*. International Journal of Forecasting, 27, 333-346.

[13] Bernanke, B. S., Boivin, J., Eliasz, P., (2005), *Measuring the Effects of Monetary Policy: a Factor-Augmented Vector Autoregressive (FAVAR) Approach*. The Quarterly Journal of Economics, 120, 387-422.

[14] Bernanke, B. S., Gertler, M., Watson, M., Sims, C. A., Friedman, B. M., (1997), *Systematic Monetary Policy and the Effects of Oil Price Shocks*. Brookings papers on economic activity, 1997(1), 91-157.

[15] Bittanti, S., (1986), *Deterministic and stochastic linear periodic systems*. Time series and linear systems (pp. 141–182). Berlin: Springer-Verlag.

[16] Bittanti, S., Colaneri, P., (2009), *Periodic systems: filtering and control*. London: Springer-Verlag.

[17] Blasque, F., Koopman, S. J., Mallee, M., (2015), *Low Frequency and Weighted Likelihood Solutions for Mixed Frequency Dynamic Factor Models*. Tinbergen Institute Discussion Papers 14-105/III, Tinbergen Institute.

[18] Boivin, J., Giannoni, M., (2006), *Has monetary policy become more effective*. Review of Economics and Statistics, vol. 88, pp. 445-462.

[19] Brown, T. M., (1960), *Simultaneous Least Squares: A Distribution Free Method of Equation System Structure Estimation*. International Economic Review, 1, 173-191.

[20] Camacho, M., (2013), *Mixed-frequency VAR models with Markov-switching dynamics*. Economics Letters, 121, 369-373.

[21] Carriero, A., Clark, T. E., Marcellino, M., (2015), *Realtime nowcasting with a Bayesian mixed frequency model with stochastic volatility*. Journal of the Royal Statistical Society Series A, 178, Part 4, 837-862.

[22] Chen, B., Zadrozny, P. A., (1998), *An Extended Yule-Walker Method for Estimating a Vector Autoregressive Model with Mixed Frequency Data*. Advances in Econometrics 13, 47-73.

[23] Castelnuovo, E., Surico, P., (2010), *Monetary policy, inflation expectations and the price puzzle*. The Economic Journal, vol. 120, pp. 1262-1283.

[24] Chen, B., Zadrozny, P. A., (1998), *An Extended Yule-Walker Method for Estimating a Vector Autoregressive Model with Mixed Frequency Data*. Advances in Econometrics 13, pp. 47-73.

[25] Chen, W., Anderson, B. D., Deistler, M., Filler, A., (2012), *Properties of Blocked Linear System*. Automatica, 48, 2520-2525.

[26] Christiano, L.J., Eichenbaum, M., (1987), *Temporal aggregation and structural inference in macroeconomics*. Carnegie-Rochester Conference Series on Public Policy, 26, pp. 64-130.

[27] Christiano, L. J., Eichenbaum, M. and Evans, C., (2005), *Nominal rigidities and the dynamic effects of a shock to monetary policy*. Journal of Political Economy, vol. 113, pp. 1-45.

[28] Cimadomo, J., D'Agostino, A., (2015), *Combining Time Variation and Mixed Frequencies: an Analysis of Government Spending Multipliers in Italy*. Journal of Applied Econometrics, DOI: 10.1002/jae.2489.

[29] Clarida, R., Galí, J., Gertler, M., (2000), *Monetary policy rules and macroeconomic stability: evidence and some theory*. Quarterly Journal of Economics, 115(1) pp. 147-180.

[30] Clark, T. E., (2011), *Real-Time Density Forecasts From Bayesian Vector Autoregressions With Stochastic Volatility*. Journal of Business & Economic Statistics, 29:3, 327-341, DOI: 10.1198/jbes.2010.09248.

[31] Clements, M., Galvão, A., (2008), Macroeconomic Forecasting With Mixed Frequency Data: Forecasting Output Growth in the United States. Journal of Business & Economic Statistics, 26, pp. 546-554.

[32] Cumby, R. E., Huizinga, J., (1992) *Testing the Autocorrelation Structure of Disturbances in Ordinary Least Squares and Instrumental Variables Regressions*. Econometrica, vol. 60(1), pp. 185-195.

[33] Del Negro, M., Schorfheide, F., (2011), *Bayesian Macroeconometrics*. The Oxford Handbook of Bayesian Econometrics, ed. by J. Geweke, G. Koop, and H. van Dijk, 293-389. Oxford University Press.

[34] Doz, C., Giannone, D., Reichlin, L., (2006), *A Two-Step Estimator for Large Approximate Dynamic Factor Models Based on Kalman Filtering.* Journal of Econometrics, vol. 164, Issue 1, 188–205.

[35] Durbin, J., Koopman, S. J., (2012), *Time Series Analysis by State Space Methods, Second Edition.* Oxford Statistical Science, 2nd edition. Oxford University Press.

[36] Fève, P., Matheron, J., Sahuc, J., (2009), *Minimum Distance Estimation and Testing of DSGE Models from Structural VARs.* Oxford Bulletin of Economics and Statistics, Department of Economics, University of Oxford, vol. 71(6), pp. 883-894.

[37] Foroni, C., Guerin, P., Marcellino, M., (2014), *Markov-switching mixed-frequency VAR models.* International Journal of Forecasting, http://dx.doi.org/10.1016/j.ijforecast.2014.05.003.

[38] Foroni, C., Marcellino, M., (2014), *Mixed Frequency Structural Models: Identification, Estimation, and Policy Analysis.* Journal of Applied Econometrics, vol. 29, Issue 7, pp. 1118-1144.

[39] Foroni, C., Marcellino, M., (2016), *Mixed Frequency Structural Vector Auto-Regressive Models.* Journal of the Royal Statistical Society Series A, Royal Statistical Society, vol. 179(2), pp. 403-425, 02.

[40] Fuhrer, J. C., Rudebusch, G. D., (2004), *Estimating the Euler equation for output.* Journal of Monetary Economics, 51(6), pp. 1133-1153.

[41] Friedlander, B., Stoica, P., Söderström, T., (1985), *Optimal Instrumental Variable Estimates of the AR Parameters of an ARMA Process.* IEEE Transactions on Automatic Control, AC-30, pp. 1066-1074.

[42] Ghysels, E., (2012), *Mixed Frequency Vector Autoregressive Models.* Mimeo.

[43] Ghysels, E., (2016), *Macroeconomics and the Reality of Mixed Frequency Data.* Journal of Econometrics, vol. 193 pp. 294-314

[44] Ghysels, E., Miller, J. I., (2015), *Testing for Cointegration with Temporally Aggregated and Mixed-Frequency Time Series.* Journal of Time Series Analysis 36: 797-816. DOI: 10.1111/jtsa.12129.

[45] Ghysels, E., Owyang, M, (2011), *Taylor Monetary Policy Rules with Financial Market Expectations.* Discussion Paper, Federal Reserve Bank of Saint-Louis and UNC.

[46] Ghysels, E., Santa-Clara, P., Valkanov, R., (2006), *Predicting volatility: getting the most out of return data sampled at different frequencies.* Journal of Econometrics, 131, pp. 59-95.

[47] Ghysels, E., Sinko, A., Valkanov, R., (2006). *MIDAS regressions: Further results and new directions.* Econometric Reviews, 26, 53-90.

[48] Giannone, D., Monti, F., Reichlin, L., (2014), *Exploiting the Monthly Dataflow in Structural Forecasting.* Working paper.

[49] Giannone, D., Reichlin, R. and Small, D. (2008), *Nowcasting GDP and Inflation: the Real-Time Informational Content of Macroeconomic Data Releases.* Journal of Monetary Economics, 55, 665-676.

[50] Giordani, P. M., Pitt, K., Kohn, R., (2011), *Bayesian Inference for Time Series State Space Models.* The Oxford Handbook of Bayesian Econometrics, ed. by J. Geweke, G. Koop, and H. K. van Dijk, 61-124. Oxford University Press.

[51] Götz, T. B., Hauzenberger, K., (2015), *Time-Varying Mixed-Frequency Vector Autoregressive Models.* Working paper.

[52] Götz, T. B., Hecq, A., Urbain. J. P., (2013), *Testing for Common Cycles in Non-stationary VARs with Varied Frequency Data.* In VAR Models in Macroeconomics – New Developments and Applications: Essays in Honor of Christopher A. Sims (Advances in Econometrics, Volume 32), Fomby TB, Kilian L, Murphy A (eds.) Emerald Group Publishing Limited: Bingley, UK, pp. 361-393.

[53] Grasselli, O. M., Longhi, S., (1988), *Zeros and Poles of Linear Periodic Multivariable Discrete-Time Systems.* Circuits Systems Signal Process, vol. 7, 3, 361-380.

[54] Guerron-Quintana, P., Inoue, A., Kilian, L., (2017), *Impulse Response Matching Estimators for DSGE Models.* Journal of Econometrics, 196, pp. 144-155.

[55] Hall, A., (1995), *Residual Autocovariances and Unit Root Tests based on Instrumental Variable Estimators from Time Series Regression Models.* Journal of Time Series Analysis, vol. 16(6), pp. 555-569.

[56] Hall, A., Inoue, A., Nason, J., Rossi, B., (2012), *Information Criteria for Impulse Response Function Matching Estimation of DSGE models.* Journal of Econometrics, Elsevier, vol. 170(2), pp 499-518.

[57] Hamilton, J. D., (1989), *A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle*. Econometrica, vol. 57, no. 2, 357-384.

[58] Hansen, L. P., Heaton, J., Yaron, A., (1996), *Finite Sample Properties of Some Alternative GMM Estimators*. Journal of Business & Economic Statistics, 14, pp. 262-80.

[59] Higham, N. J., (2008), *Functions of Matrices: Theory and Computation*. Philadelphia: Society for Industrial and Applied Mathematics.

[60] de Jong, P., (1989), *Smoothing and interpolation with the state-space model*. Journal of the American Statistical Association, vol. 84, 1085–1088.

[61] Kim, T. B., (2010), *Temporal Aggregation Bias and Mixed Frequency Estimation of New Keynesian Model*. Mimeo.

[62] Kim, T. B., (2010), *Temporal Aggregation Bias and Mixed Frequency Estimation of New Keynesian Model*. Mimeo.

[63] Klein, R., Vella, F., (2006), *Estimating a Class of Triangular Simultaneous Equation Models without Exclusion Restrictions*. IZA Discussion Paper 2378.

[64] Koelbl, L., Braumann, A., Felsenstein, E., Deistler, M., (2016), *Estimation of VAR Systems from Mixed-Frequency Data: The Stock and the Flow Case* . Dynamic Factor Models in Advances in Econometrics, Emerald Group Publishing Limited vol. 35, pp. 43-73.

[65] Kuzin, V., Marcellino, M., Schumacher, C., (2011), *Midas vs. Mixed Frequency VAR: Nowcasting GDP in the Euro Area*. International Journal of Forecasting, 27, pp. 529-542.

[66] Kuzin, V., Marcellino, M., Schumacher, C., (2013), *Pooling Versus Model Selection for Nowcasting GDP with Many Predictors: Empirical Evidence for Six Industrialized Countries*. Journal of Applied Econometrics, 28, 392-411.

[67] Lanne, M., Lütkepohl, H., (2008), *Identifying monetary policy shocks via changes in volatility*. Journal of Money, Credit and Banking, Vol. 40, pp. 1131-1149.

[68] Lütkepohl, H., (2012), *New Introduction to Multiple Time Series Analysis (2nd edn.)*. Springer-Verlag Berlin.

[69] Lütkepohl, H., (1987), *Forecasting Aggregated Vector ARMA Processes*. Berlin: Springer-Verlag.

[70] Malinvaud, E., (1970), *Statistical Methods of Econometrics. (2nd edn.)*. Amsterdam: North-Holland Publishing Co.

[71] Marcellino, M., (1999), *Some Consequences of Temporal Aggregation in Empirical Analysis*, Journal of Business & Economic Statistics, 17:1, pp. 129-136.

[72] Marcellino, M., Porqueddu, M., Venditti, F., (2015), *Short-term GDP forecasting with a mixed frequency dynamic factor model with stochastic volatility*. Journal of Business & Economic Statistics, DOI: 10.1080/07350015.2015.1006773.

[73] Marcellino, M., Sivec, V., (2015), *Monetary, Fiscal and Oil Shocks: Evidence based on Mixed Frequency Structural FAVARs*. CEPR Discussion Papers 10610, C.E.P.R. Discussion Papers.

[74] Mariano, R. S., Murasawa, Y., (2003). *A new coincident index of business cycles based on monthly and quarterly series*. Journal of Applied Econometrics, 18(4), pp. 427-443.

[75] Mariano, R. S., Murasawa, Y., (2010), *A Coincident Index, Common Factors, and Monthly Real GDP*. Oxford Bulletin of Economics and Statistics, 72(1), pp. 27-46.

[76] Nakamura, M., (1960), *A Note on the Consistency of Simultaneous Least Squares Estimation*. International Economic Review, vol. 1(3), pp. 192-197.

[77] Newey, W. K., McFadden, D., (1994), *Large Sample Estimation and Hypothesis Testing*. In The Handbook of Econometrics, edited by Robert F. Engle and Daniel McFadden, vol. 4, pp. 2111-245. Amsterdam: North-Holland.

[78] Newey, W. K., West, K. D., (1987), *A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix*. Econometrica, vol. 55(3), pp. 703–708.

[79] Pierse, R. G., Snell, A. J., (1995), *Temporal Aggregation and the Power of Tests for a Unit Root*. Journal of Econometrics, 65, pp. 333-345.

[80] Ramey, V. A., (2011), *Identifying Government Spending Shocks: It's All in the Timing*. The Quarterly Journal of Economics, 126(1), 1-50.

[81] Rodriguez, A., Puggioni, G., (2010), *Mixed frequency models: Bayesian approaches to estimation and prediction*. International Journal of Forecasting, 26, 293-311.

[82] Rigobon, R., (2003) *Identification through Heteroskedasticity.* Review of Economics and Statistics, 85, pp. 777-92.

[83] Rossana, R. J., and Seater, J. J., (1992), *Aggregation, Unit Roots and the Time Series Structure of Manufacturing Real Wages.* International Economic Review, 33, pp. 159-179.

[84] Rotemberg, J. J., Woodford, M., (1997), *An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy.* NBER Macroeconomics Annual 1997, pp. 297-346.

[85] Schorfheide, F. and Song, D. (2015), *Real-time Forecasting with a Mixed Frequency VAR.* Journal of Business & Economic Statistics, Volume 33, Issue 3, p.366-380.

[86] Seong, B., Ahn, S.K., Zadrozny P. A., (2013), *Cointegration Analysis with Mixed-Frequency Data.* Journal of Time Series Analysis, 34, 194-205.

[87] Wei, W. W. S., (1981), *Effect of Systematic Sampling on ARIMA Models.* Communication in Statistical-Theoretical Mathematics, 10, pp. 1389-2398.

[88] Weiss, A. A., (1984), *Systematic Sampling and Temporal Aggregation in Time Series Models.* Journal of Econometrics, pp. 271-281.

[89] Zadrozny, P. A., (1988), *Gaussian Likelihood of Continuous-Time ARMAX Models When Data Are Stocks and Flows at Different Frequencies.* Econometric Theory, Cambridge University Press, vol. 4(01), 108-124.

[90] Zamani, M., Chen, W., Anderson, B. D., Deistler, M., Filler, A., (2011), *On the Zeros of Blocked Linear Systems with Single and Mixed Frequency Data.* 50th IEEE Conference on Decision and Control, and European Control Conference (CDC-ECC), Orlando, FL, USA, December 12-15.