

Testing for common cycles in non-stationary VARs with varied frequency data

Citation for published version (APA):

Götz, T. B., Hecq, A. W., & Urbain, J. R. Y. J. (2013). Testing for common cycles in non-stationary VARs with varied frequency data. (GSBE Research Memorandum; No. 002). Maastricht: GSBE.

Document status and date:

Published: 01/01/2013

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Thomas B. Götz, Alain Hecq,
Jean-Pierre Urbain

**Testing for Common Cycles in
Non-Stationary VARs with
Varied Frequency Data**

RM/13/002

GSBE

Maastricht University School of Business and Economics
Graduate School of Business and Economics

P.O. Box 616
NL- 6200 MD Maastricht
The Netherlands

TESTING FOR COMMON CYCLES IN NON-STATIONARY VARs WITH VARIED FREQUENCY DATA*

Thomas B. Götz[†] Alain Hecq
Jean-Pierre Urbain

Maastricht University, SBE, Department of Quantitative Economics

January 17, 2013

Abstract

This paper proposes a new way for detecting the presence of common cyclical features when several time series are observed/sampled at different frequencies, hence generalizing the common-frequency approach introduced by Engle and Kozicki (1993) and Vahid and Engle (1993). We start with the mixed-frequency VAR representation investigated in Ghysels (2012) for stationary time series. For non-stationary time series in levels, we show that one has to account for the presence of two sets of long-run relationships. The first set is implied by identities stemming from the fact that the differences of the high-frequency $I(1)$ regressors are stationary. The second set comes from possible additional long-run relationships between one of the high-frequency series and the low-frequency variables. Our transformed VECM representations extend the results of Ghysels (2012) and are very important for determining the correct set of variables to be used in a subsequent common cycle investigation. This has some empirical implications both for the behavior of the test statistics as well as for forecasting. Empirical analyses with the quarterly real GNP and monthly industrial production indices for, respectively, the U.S. and Germany illustrate our new approach. This is also investigated in a Monte Carlo study, where we compare our proposed mixed-frequency models with models stemming from classical temporal aggregation methods.

JEL Codes: C32

JEL Keywords: Common Cycles, Mixed Frequency VAR, cointegration, MIDAS

*We thank Lutz Kilian and the participants of the 32nd Annual International Symposium on Forecasting in Boston and of the 12th Annual Advances in Econometrics Conference in Dallas for useful suggestions and comments on earlier versions of the paper.

[†]Corresponding author: Thomas Götz, Department of Quantitative Economics, School of Business and Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands. Email: t.gotz@maastrichtuniversity.nl, Tel.: +31 43 388 3578, Fax: +31 43 388 2000.

1 Introduction

Time series are collected and observed at different sampling frequencies. For instance, national account variables such as the gross domestic product (GDP) are often available on a quarterly basis whilst many other indicators such as the industrial production index, the unemployment rate or price indices are supplied at a monthly frequency. Several variables are only reported annually (e.g. CO2 emission, health outlook, etc.) whilst most financial time series (stock prices, interest rates, exchange rates, etc.) are recorded almost continuously and are easily and at least freely obtained at a daily frequency. In this situation it is common in macroeconomics to focus on the lower frequency unless a good proxy for the variable of interest is available at higher frequency. For instance, if one wants to estimate a vector autoregressive model (VAR) with GDP included, one would probably run a VAR model on quarterly time series. The higher frequency variables (monthly, daily,...) are either averaged over quarters (usually for flow variables) or matched at the end of each quarter (usually for stock variables).

More recently, researchers have developed techniques for modeling, in a unified framework, variables sampled at different frequencies (see notably the MIDAS approach in Andreou, Ghysels and Kourtellis, 2010). Most of the mixed-frequency literature which relies on what Ghysels (2012) calls observation-driven models¹ (as introduced in Cox, 1981) is concerned with a regression framework where the dependent variable is the low-frequency variable and the high-frequency variables are in the regressors. Only few papers have extended the analysis towards a multivariate setting, examples are Qian (2010) and Ghysels (2012). We refer the reader to Forni, Ghysels and Marcellino (2012) for a current survey on this topic.

Qian (2010) deals with a mixed-frequency VAR in which some variables are possibly unobserved to account for the fact that some series may be observed at lower frequencies. Ghysels (2012), in contrast, stacks the high- and low-frequency variables into a vector and formulates a VAR in terms of this vector. Interestingly, the model does not involve any latent variables, and thereby no latent shocks, but is formulated exclusively in terms of observable data explaining the term "observation-driven model". Note that due to its observation-driven nature, the model's related impulse response functions are entirely driven by observable shocks. As pointed out by Forni et al. (2012) the mixed-frequency VAR à la Ghysels (2012) jointly specifies the dynamics of all variables concerned without imposing any a-priori restrictions. Furthermore, its ability to apply standard tools of the VAR literature, e.g. impulse response functions, to a mixed-frequency setting gives it a significant edge over univariate approaches such as the aforementioned MIDAS approach.

Note that the setup proposed in Ghysels (2012) acts as a starting point for this work, implying that we deal with observation-driven instead of parameter-driven models.²However,

¹In contrast to the parameter-driven models such as state space models à la Zadzorny (1988). See Seong, Ahn and Zadzorny (2012) for VECMs in state-space forms.

²Forni et al. (2012) stress that only under the ideal circumstances of ignored parameter estimation errors and assumed correct specification of the model are the state-space models unrivalled. These ideal conditions are often, however, difficult to meet. As an example, correct specification of the model in high-frequency is unusually complex due to the missing observations in the regressand. Furthermore, properly specifying some

it exclusively copes with covariance-stationary variables whilst many macroeconomic variables are non-stationary and possibly cointegrated. Hence, an extension of a mixed-frequency VAR model along these lines is of important relevance. We show that for non-stationary time series in levels, one has to account for the presence of two sets of long-run relationships. The first set is characterized by identities stemming from the fact that the differences of the high-frequency $I(1)$ regressors are stationary. The second set comes from additional long-run relationships between one of the high-frequency series and the low-frequency variables. Without this latter set of long-run relationships, the VAR in first differences looks very much like in Ghysels (2012). However, we point out additional zero-restrictions that are relevant for our new common cyclical feature framework. These transformed VECM representations allow us to extend the common cyclical feature analysis (see Engle and Kozicki, 1993, or Vahid and Engle, 1993) to a mixed-frequency setup. As such, these mixed-frequency VARs present an alternative to temporally aggregating the high-frequency variable in order to obtain a VAR in a common frequency (Marcellino, 1999, or Hecq, 1998).

The rest of the paper is organized as follows. We start with the stacked VAR representation proposed in Ghysels (2012) and show how to extend it to account for $I(1)$, possibly cointegrated, variables. In Section 3 we propose to formalize an alternative VAR representation proving useful for testing for the presence of common cyclical features. A corresponding likelihood ratio type test, for which the choice of which dynamic variables to include has been carefully discussed in Section 3, is proposed in Section 4. Notice that our new approach is set up in a maximum likelihood framework such that it is best for a small number of variables and for a relatively small observation frequency. In Section 5, empirical analyses with the quarterly GNP and monthly IPIs for, respectively, the U.S. and Germany are presented to illustrate our approach. Section 5 also contains a small forecasting exercise emphasizing its potential usefulness over classical temporal aggregation methods. Section 6 presents a Monte Carlo study investigating the frequencies with which information criteria detect different lag lengths and rejection frequencies of the aforementioned likelihood ratio type test for both, our proposed mixed-frequency and temporal aggregation based models. Section 7 concludes.

2 A mixed-frequency VECM

Let us start from a two variable mixed-frequency system³, where y_t , $t = 1 \dots T$ is the low-frequency variable and $x_{t-i/m}^{(m)}$ the high-frequency variables with m high-frequency observations per low-frequency period t . In a year/month-example $m = 12$ and the value of i indicates the specific month under consideration, ranging from January ($x_{t-11/m}^{(m)}$) until December ($x_t^{(m)}$). Obviously, $x_{t-m/m}^{(m)} = x_{t-1}^{(m)}$. Note that as far as lag operators, L , or difference operators, Δ ,

high-frequency variables may be so demanding that related "specification errors in one equation spread to the entire system" (Foroni et al., 2012). As a consequence, this paper does not concern itself with state-space forms or Kalman filtering, a formal comparison of such parameter-driven models to our approach is left for future research.

³An extension to higher dimensional multivariate systems is straightforward.

are concerned, we make a distinction between high- and low-frequency operators. To be more precise, L denotes the low-frequency lag operator, i.e., $Ly_t = y_{t-1}$ or $Lx_{t-i/m}^{(m)} = x_{t-1-i/m}^{(m)}$, whereas L_m denotes the high-frequency counterpart which is only applicable to the high-frequency variables, i.e., $L_mx_{t-i/m}^{(m)} = x_{t-i/m-1/m}^{(m)} = x_{t-(i+1)/m}^{(m)}$. Similar rules apply for the difference operators, Δ and Δ_m . Finally, note that $L_mx_{t-(m-1)/m}^{(m)} = x_{t-1}^{(m)}$ and, by the same logic, $\Delta_mx_{t-(m-1)/m}^{(m)} = x_{t-(m-1)/m}^{(m)} - x_{t-1}^{(m)}$. These notational conventions are similar to the ones used in Clements and Galvão (2007, 2009) or Götz, Hecq and Urbain (2012a, 2012b). The following table illustrates the notation for a year/month-example:

Notation	$t = 2011, m = 12$
$x_{t+1-(m-1)/m}^{(m)} = x_{t+1/12}^{(12)}$	$x_{2012,Jan}^{(12)}$
$x_t^{(m)}$	$x_{2011,Dec}^{(12)}$
$x_{t-1/m}^{(m)}$	$x_{2011,Nov}^{(12)}$
\vdots	\vdots
$x_{t-(m-1)/m}^{(m)} = x_{t-11/12}^{(12)}$	$x_{2011,Jan}^{(12)}$
$x_{t-m/m}^{(m)} = L_mx_{t-11/12}^{(12)} = x_{t-1}^{(12)}$	$x_{2010,Dec}^{(12)}$
$x_{t-1-1/m}^{(m)}$	$x_{2010,Nov}^{(12)}$

Considering each high-frequency variable such that $X_t^{(m)} = (x_t^{(m)}, x_{t-1/m}^{(m)}, \dots, x_{t-(m-1)/m}^{(m)})'$, Ghysels (2012) proposes to consider, after achieving stationarity of I(1) series (e.g. quarterly growth rate of output and the monthly growth rate of the production index), the VAR(p):

$$Z_t = \Gamma_1 Z_{t-1} + \dots + \Gamma_p Z_{t-p} + \varepsilon_t, \quad (1)$$

where $Z_t = (y_t, X_t^{(m)})'$ and $\varepsilon_t \sim \text{i.i.d. } N(\mathbf{0}, I_{m+1})$.⁴ In other words, the high-frequency observations per low-frequency period are treated as different low-frequency variables and stacked in a vector, together with the low-frequency variable y_t , to form a VAR. This particular composition implies that our work is related to a periodic framework introduced by Gladyshev (1961). Note that compared to Ghysels (2012) we place the low-frequency variable first to facilitate the interpretation of the presence of common cycles we develop in Section 4.

Equation (1) is easy to estimate on small systems but relies on the assumption that the time series are genuinely stationary. Now, assuming that series in Z_t are I(1) and that there exists some cointegration between the variables, (1) in first differences is not correctly specified. This can be seen by computing the VECM representation of (1) when both the high- and low-frequency variables are $I(1)$ such that

$$\Delta Z_t = \tilde{\Gamma}_1 \Delta Z_{t-1} + \dots + \tilde{\Gamma}_{p-1} \Delta Z_{t-p+1} + \Pi Z_{t-1} + \varepsilon_t, \quad (2)$$

⁴It is possible to include deterministic components such as a constant, trends or seasonal dummies here as well. For illustrative purposes we skip it in the theoretical sections of the paper. In the application and Monte Carlo study, however, a vector of intercepts is included.

with $\tilde{\Gamma}_i = -\sum_{k=i+1}^p \Gamma_k$, $i = 1, \dots, p-1$ and $\Pi = -(I - \sum_{j=1}^p \Gamma_j) = \alpha\beta'$ with $\text{rank}(\Pi) = (r_0 + r_1) < m + 1$.

Let us focus on the rank of Π in the presence and absence of cointegration between the low- and high-frequency variables. r_0 denotes a first set of known or prespecified cointegrating vectors stemming from the fact that first differences of the high-frequency $I(1)$ variables are stationary. r_1 will denote the potential additional long-run relationship between the two different series. Note that in case of cointegration it is sufficient to consider only one relationship between y_t and one of the various $x_{t-i/m}^{(m)}$'s. The reason is that if cointegration is present between the two variables, y is cointegrated with each of high-frequency x 's falling into the respective t -period. Hence, it does not matter which of these high-frequency candidates is used to capture the cointegrating relationship as long as one is used (see Götz et al., 2012a, for details). For simplicity, we model cointegration employing the end-of-period observation of the high-frequency variable, i.e., $x_t^{(m)}$.

Anticipating the results of the empirical section, while r_0 and the associated cointegrating vectors are known, a test on r_1 will be determined using the approach of Horvath and Watson (1995) for testing for cointegration when some of the cointegrating vectors are prespecified.⁵ In particular, we will test the null hypothesis $H_0 : \text{rank}(\Pi) = r_0$ against $H_A : \text{rank}(\Pi) = r_0 + r_1$, where the cointegrating vectors imposed under the null are precisely the prespecified ones discussed above. The test itself is a likelihood ratio test, where, under the alternative model, additional unknown cointegrating vectors, the parameters of which being chosen so as to maximize the log-likelihood function, are added to the identities. In our case, only one unknown cointegrating vector to capture potential cointegration between the low- and high-frequency variables is added and is of the form $(1 : -\theta : 0_{1 \times (m-1)})$, where the coefficient of y is normalized to one and θ is the cointegrating coefficient.

Our first results can be summarized in the following lemma and example.

Observation 1 *Assume that the high-frequency variables, $x_{t-i/m}^{(m)}$, and the low-frequency variable, y_t , are $I(1)$. In case of no cointegration between the low- and high-frequency variables, $\text{rank}(\Pi) = r_0$. $r_0 = m - 1$ in a bivariate analysis. In case of cointegration between the low- and high-frequency variables, $\text{rank}(\Pi) = r_0 + r_1$. $r_1 = 1$ in a bivariate analysis.*

Trivially, each high-frequency variable in the subvector X_t of Z_t is cointegrated with the other high-frequency ones. Since there are m such variables in Z_t in a bivariate case, we have $m - 1$ prespecified cointegrating relationships r_0 . We model any additional cointegrating relationships between the low- and high-frequency variables by considering one relationship between y_t and $x_t^{(m)}$ as argued before. This adds $r_1 = 1$ to $\text{rank}(\Pi)$ in a bivariate example. These arguments can straightforwardly be extended to a higher dimensional framework.

⁵Horvath and Watson (1995) show that accounting for the fact that some cointegrating vectors are prespecified leads to improved power results in terms of testing for potential further, unknown, cointegrating relationships.

Example 2 Let us consider a very simple example where the low- and high-frequencies are assumed to be quarterly and monthly, respectively, and where y_t and the elements of $X_t^{(m)}$ are defined as

$$\begin{aligned} y_t &= \theta x_t^{(3)} + \xi_t \text{ with } \xi_t = \rho \xi_{t-1} + \nu_t, \text{ where } |\rho| < 1 \text{ and } \nu_t \sim \text{i.i.d. } N(0, 1) \\ x_{t-i/3}^{(3)} &= x_{t-(i+1)/3}^{(3)} + u_{t-i/3}^{(3)} \text{ for } i = 0, 1, 2 \text{ and where all } u_{t-i/3}^{(3)} \sim \text{i.i.d. } N(0, 1), \end{aligned}$$

implying that cointegration between the low- and high-frequency variables is present. Representing this in a VAR framework as explained above yields

$$\begin{pmatrix} y_t \\ x_t^{(3)} \\ x_{t-1/3}^{(3)} \\ x_{t-2/3}^{(3)} \end{pmatrix} = \begin{pmatrix} \rho & (1-\rho)\theta & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1}^{(3)} \\ x_{t-1-1/3}^{(3)} \\ x_{t-1-2/3}^{(3)} \end{pmatrix} + \begin{pmatrix} \theta \sum_{i=0}^2 u_{t-i/3}^{(3)} + \nu_t \\ \sum_{i=0}^2 u_{t-i/3}^{(3)} \\ u_{t-1/3}^{(3)} + u_{t-2/3}^{(3)} \\ u_{t-2/3}^{(3)} \end{pmatrix}.$$

The corresponding VECM representation looks as follows:

$$\begin{pmatrix} \Delta y_t \\ \Delta x_t^{(3)} \\ \Delta x_{t-1/3}^{(3)} \\ \Delta x_{t-2/3}^{(3)} \end{pmatrix} = \begin{pmatrix} \rho - 1 & (1-\rho)\theta & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1}^{(3)} \\ x_{t-1-1/3}^{(3)} \\ x_{t-1-2/3}^{(3)} \end{pmatrix} + \begin{pmatrix} \theta \sum_{i=0}^2 u_{t-i/3}^{(3)} + \nu_t \\ \sum_{i=0}^2 u_{t-i/3}^{(3)} \\ u_{t-1/3}^{(3)} + u_{t-2/3}^{(3)} \\ u_{t-2/3}^{(3)} \end{pmatrix},$$

such that $\text{rank}(\Pi) = r_0 + r_1 = 2 + 1 = 3$ and with long-run relationships for instance

$$\beta' Z_{t-1} = \begin{pmatrix} \rho - 1 & -(\rho - 1)\theta & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1}^{(3)} \\ x_{t-1-1/3}^{(3)} \\ x_{t-1-2/3}^{(3)} \end{pmatrix}.$$

If, however, no cointegration exists between the low- and high-frequency variables, e.g. both variables follow random walks,

$$\begin{aligned} y_t &= y_{t-1} + \nu_t \\ x_{t-i/3}^{(3)} &= x_{t-(i+1)/3}^{(3)} + u_{t-i/3}^{(3)} \text{ for } i = 0, 1, 2, \end{aligned}$$

where ν_t and the $u_{t-i/3}^{(3)}$'s are i.i.d. standard normally distributed, we get the following VECM representation:

$$\begin{pmatrix} \Delta y_t \\ \Delta x_t^{(3)} \\ \Delta x_{t-1/3}^{(3)} \\ \Delta x_{t-2/3}^{(3)} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1}^{(3)} \\ x_{t-1-1/3}^{(3)} \\ x_{t-1-2/3}^{(3)} \end{pmatrix} + \begin{pmatrix} \nu_t \\ \sum_{i=0}^2 u_{t-i/3}^{(3)} \\ u_{t-1/3}^{(3)} + u_{t-2/3}^{(3)} \\ u_{t-2/3}^{(3)} \end{pmatrix},$$

such that $\text{rank}(\Pi) = r_0 = m - 1 = 2$ and the (prespecified) long-run cointegrating relationships are

$$\beta' Z_{t-1} = \begin{pmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1}^{(3)} \\ x_{t-1-1/3}^{(3)} \\ x_{t-1-2/3}^{(3)} \end{pmatrix}.$$

3 An alternative "desirable" VAR/VECM representation

The previous example illustrates that, starting from the VAR(p) model in (1), we can compute its VECM representation in the standard way and get (2). However, we end up with low-frequency differences of all variables, including the high-frequency ones. For instance, we would have a VECM between the quarterly growth rate of the GDP and the quarterly growth rate of the industrial production index. This is not the series that Ghysels (2012) would use in his investigation of a mixed-frequency VAR neither the ones we would like to consider in our study.

Given that in a unified framework, mixed-frequency error-correction models are characterized by low- and high-frequency differences for the low- and high-frequency variables, respectively (Götz et al., 2012a, or Miller, 2011), it is important to transform the VECM accordingly.

We compute the transformed VAR (see Warne, 1993) as e.g. illustrated in Hecq and Jacobs (2009). In particular, we start from the VECM in (2) where $\Pi = \alpha\beta'$ such that α and β are $(m+1) \times r$ matrices with $r = r_0 + r_1$ being the number of cointegrating relationships. To obtain the transformed VAR representation, we introduce the $(m+1) \times (m+1)$ matrix $C = (M_y : \beta : M_x)'$, with selection matrices $M_y = (1 \ \mathbf{0}_{1 \times m})' \mathbb{1}_{(r=m-1)}$ and $M_x = (\mathbf{0}_{1 \times m} \ 1)'$. The indicator function in M_y makes sure we add a corresponding selection row in case y and x are not cointegrated. Premultiplying both sides of (2) by C gives

$$\begin{aligned} C\Delta Z_t &= C\tilde{\Gamma}_1 C^{-1} C\Delta Z_{t-1} + \dots + C\tilde{\Gamma}_{p-1} C^{-1} C\Delta Z_{t-p+1} + C\Pi C^{-1} CZ_{t-1} + C\varepsilon_t \\ \Leftrightarrow CZ_t &= (I + C\tilde{\Gamma}_1 C^{-1} + C\Pi C^{-1}) CZ_{t-1} + (C\tilde{\Gamma}_2 C^{-1} - C\tilde{\Gamma}_1 C^{-1}) CZ_{t-2} + \\ &\quad \dots + (-C\tilde{\Gamma}_{p-1} C^{-1}) CZ_{t-p} + C\varepsilon_t. \end{aligned} \tag{3}$$

Whether or not there is cointegration between y and x , it is possible to rewrite (3) such that we finally obtain

$$\Delta Z_t^* = \Phi_1 \Delta Z_{t-1}^* + \Phi_2 \Delta Z_{t-2}^* + \dots + \Phi_p \tilde{Z}_{t-p} + \tilde{\varepsilon}_t, \tag{4}$$

where

$$\Delta Z_t^* = \begin{pmatrix} \Delta y_t \\ \Delta_m x_t^{(m)} \\ \vdots \\ \Delta_m x_{t-(m-2)/m}^{(m)} \\ \Delta_m x_{t-(m-1)/m}^{(m)} \end{pmatrix}, \quad \tilde{Z}_t = \begin{pmatrix} y_t \\ \Delta_m x_t^{(m)} \\ \vdots \\ \Delta_m x_{t-(m-2)/m}^{(m)} \\ x_{t-(m-1)/m}^{(m)} \end{pmatrix} \quad \text{and} \quad \tilde{\varepsilon}_t = \begin{pmatrix} \epsilon_{t,y} \\ \Delta_m \epsilon_t^{(m)} \\ \vdots \\ \Delta_m \epsilon_{t-(m-2)/m}^{(m)} \\ \epsilon_{t-(m-1)/m}^{(m)} \end{pmatrix}.$$

In particular we can write (4) as

$$\Delta Z_t^* = \Phi_1 \Delta Z_{t-1}^* + \Phi_2 \Delta Z_{t-2}^* + \dots + \Phi_p^0 \tilde{Z}_{t-p} + \Phi_p^1 \tilde{Z}_{t-p} + \tilde{\varepsilon}_t, \quad (5)$$

where Φ_p^0 is the matrix with the central part of Φ_p and with zero-coefficients on the first and the last column. Then $\Phi_p^0 \tilde{Z}_{t-p} = \Phi_p^0 \Delta Z_{t-p}^*$ only has coefficients for $\Delta_m x_{t-p,m}, \dots, \Delta_m x_{t-p,2}$ such that the p^{th} lags of y_t and $x_{t,1}$ are not present. In the presence of additional (unknown) r_1 cointegrating vectors $\Phi_p^1 = \Phi_p - \Phi_p^0$ is of rank $r_1 > 0$. In the absence of such cointegrating vectors $\Phi_p = \Phi_p^0$ such that $\Phi_p^1 = \mathbf{0}_{(m+1) \times (m+1)}$.

Remark 3 *As far as the error term is concerned, it was necessary to assume that ε_t is i.i.d. and standard normally distributed in order for (2) not to be misspecified. Premultiplying ε_t by C in \mathfrak{B} , however, causes $\tilde{\varepsilon}_t$ to contain differences of $\epsilon_t^{(m)}, \dots, \epsilon_{t-(m-2)/m}^{(m)}$. Note, however, that these are high-frequency differences, Δ_m , implying that all terms contained in $\tilde{\varepsilon}_t$ correspond to observations within one low-frequency period t . Hence, as the transformed VAR, (5), is written in terms of the low frequency, premultiplying ε_t by C does not lead to a misspecification of the model.*

Example 4 *Let us again assume that the high- and low-frequency variables are monthly and quarterly ones, respectively. Furthermore, suppose first that y and x are not cointegrated and let us consider a VAR(1), $Z_t = \Gamma_1 Z_{t-1} + \varepsilon_t$ (without deterministic components for simplicity), from which we get to a low-frequency VECM representation, $\Delta Z_t = \alpha \beta^l Z_{t-1} + \varepsilon_t$ as:*

$$\Leftrightarrow \begin{pmatrix} \Delta y_t \\ \Delta x_t^{(3)} \\ \Delta x_{t-1/3}^{(3)} \\ \Delta x_{t-2/3}^{(3)} \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \\ \alpha_{41} & \alpha_{42} \end{pmatrix} \begin{pmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1}^{(3)} \\ x_{t-1-1/3}^{(3)} \\ x_{t-1-2/3}^{(3)} \end{pmatrix} + \begin{pmatrix} \epsilon_{t,y}^{(3)} \\ \epsilon_{t,x}^{(3)} \\ \epsilon_{t-1/3}^{(3)} \\ \epsilon_{t-2/3}^{(3)} \end{pmatrix}.$$

Now, we premultiply both sides of the equation by

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

to obtain

$$\begin{aligned} C \Delta Z_t &= C \Pi Z_{t-1} + C \varepsilon_t \\ \Leftrightarrow C Z_t &= C Z_{t-1} + C \Pi C^{-1} C Z_{t-1} + C \varepsilon_t \\ \Leftrightarrow \tilde{Z}_t &= (I + C \Pi C^{-1}) \tilde{Z}_{t-1} + \tilde{\varepsilon}_t. \end{aligned}$$

Note that

$$\tilde{Z}_t = \begin{pmatrix} y_t \\ \Delta_m x_t^{(3)} \\ \Delta_m x_{t-1/3}^{(3)} \\ x_{t-2/3}^{(3)} \end{pmatrix} = \begin{pmatrix} \Delta y_t \\ \Delta_m x_t^{(3)} \\ \Delta_m x_{t-1/3}^{(3)} \\ \Delta_m x_{t-2/3}^{(3)} \end{pmatrix} + \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix}}_A \begin{pmatrix} y_{t-1} \\ \Delta_m x_{t-1}^{(3)} \\ \Delta_m x_{t-1-1/3}^{(3)} \\ x_{t-1-2/3}^{(3)} \end{pmatrix} = \Delta Z_t^* + A \tilde{Z}_{t-1}.$$

Hence,

$$\begin{aligned}
\tilde{Z}_t &= (I + C\Pi C^{-1})\tilde{Z}_{t-1} + \tilde{\varepsilon}_t \\
\Leftrightarrow \Delta Z_t^* &= \underbrace{(I - A + C\Pi C^{-1})}_{\Phi_1 = \Phi_p^0} \tilde{Z}_{t-1} + \tilde{\varepsilon}_t \\
\Leftrightarrow \begin{pmatrix} \Delta y_t \\ \Delta_m x_t^{(3)} \\ \Delta_m x_{t-1/3}^{(3)} \\ \Delta_m x_{t-2/3}^{(3)} \end{pmatrix} &= \underbrace{\begin{pmatrix} 0 & \alpha_{11} & \alpha_{12} & 0 \\ 0 & 1 + \alpha_{21} - \alpha_{31} & \alpha_{22} - \alpha_{32} & 0 \\ 0 & \alpha_{31} - \alpha_{41} & 1 + \alpha_{32} - \alpha_{42} & 0 \\ 0 & \alpha_{41} - 1 & \alpha_{42} - 1 & 0 \end{pmatrix}}_{\Phi_1} \begin{pmatrix} y_{t-1} \\ \Delta_m x_{t-1}^{(3)} \\ \Delta_m x_{t-1-1/3}^{(3)} \\ x_{t-1-2/3}^{(3)} \end{pmatrix} + \begin{pmatrix} \epsilon_{t,y} \\ \Delta_m \epsilon_t^{(3)} \\ \Delta_m \epsilon_{t-1/3}^{(3)} \\ \epsilon_{t-2/3}^{(3)} \end{pmatrix}.
\end{aligned}$$

This shows that in the presence of prespecified cointegrating vectors only, the mixed-frequency stationary VAR specification of Ghysels (2012) does not impose the full set of restrictions on the dynamics. While consistency of the estimators is still achieved, efficiency might be affected when the VAR is not restricted accordingly. In the framework of this paper, this changes the set of variables that has to be used in the subsequent common feature analysis.

Example 5 More importantly now, if cointegration between y and x is present as in

$$\Pi = \alpha\beta' = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \\ \alpha_{41} & \alpha_{42} & \alpha_{43} \end{pmatrix} \begin{pmatrix} 1 & -\theta & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix},$$

we select $C = (\beta : M_x)'$ with $m = 3$. Note that $CZ_t \neq \tilde{Z}_t$, a problem we get around by observing that $A_\theta CZ_t = \tilde{Z}_t$, where

$$A_\theta = \begin{pmatrix} 1 & \theta & \theta & \theta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Hence, we have the following:

$$\begin{aligned}
CZ_t &= (I + C\Pi C^{-1})CZ_{t-1} + C\varepsilon_t \\
\Leftrightarrow A_\theta^{-1}\tilde{Z}_t &= (I + C\Pi C^{-1})A_\theta^{-1}\tilde{Z}_{t-1} + A_\theta^{-1}\tilde{\varepsilon}_t \\
\Leftrightarrow \tilde{Z}_t &= A_\theta(I + C\Pi C^{-1})A_\theta^{-1}\tilde{Z}_{t-1} + \tilde{\varepsilon}_t \\
\Leftrightarrow \Delta Z_t^* &= \underbrace{(A_\theta(I + C\Pi C^{-1})A_\theta^{-1} - A)}_{\Phi_1} \tilde{Z}_{t-1} + \tilde{\varepsilon}_t \\
&= (I - A + A_\theta C\Pi C^{-1} A_\theta^{-1}) \tilde{Z}_{t-1} + \tilde{\varepsilon}_t,
\end{aligned}$$

where

$$\Phi_1 = \begin{pmatrix} \alpha_{11} & \alpha_{12} - \alpha_{11} & \alpha_{13} - \theta\alpha_{11} & -\theta\alpha_{11} \\ \alpha_{21} - \alpha_{31} & 1 + \alpha_{22} - \theta\alpha_{21} + \theta\alpha_{31} & \alpha_{23} - \alpha_{33} - \theta\alpha_{21} + \theta\alpha_{31} & -\theta(\alpha_{21} - \alpha_{31}) \\ \alpha_{31} - \alpha_{41} & \alpha_{32} - \alpha_{42} - \theta\alpha_{31} + \theta\alpha_{41} & 1 + \alpha_{33} - \alpha_{43} - \theta\alpha_{31} + \theta\alpha_{41} & -\theta(\alpha_{31} - \alpha_{41}) \\ \alpha_{41} & \alpha_{42} - \theta\alpha_{41} - 1 & \alpha_{43} - \theta\alpha_{41} - 1 & -\theta\alpha_{41} \end{pmatrix}.$$

Note that the first and fourth columns of Φ_1 are linearly dependent, i.e., $\Phi_1 e'_1 = -\theta \Phi_1 e'_4$, where e_i is a row-vector of size $m+1$ with a value of one at the i^{th} position and zeroes everywhere else. Hence, we can write $\Delta Z_t^* = \Phi_1^0 \tilde{Z}_{t-1} + \Phi_1^1 \tilde{Z}_{t-1} + \tilde{\varepsilon}_t$, where

$$\Phi_1^0 = \begin{pmatrix} 0 & \alpha_{12} - \alpha_{11} & \alpha_{13} - \theta\alpha_{11} & 0 \\ 0 & 1 + \alpha_{22} - \alpha_{32} - \theta\alpha_{21} + \theta\alpha_{31} & \alpha_{23} - \alpha_{33} - \theta\alpha_{21} + \theta\alpha_{31} & 0 \\ 0 & \alpha_{32} - \alpha_{42} - \theta\alpha_{31} + \theta\alpha_{41} & 1 + \alpha_{33} - \alpha_{43} - \theta\alpha_{31} + \theta\alpha_{41} & 0 \\ 0 & \alpha_{42} - \theta\alpha_{41} - 1 & \alpha_{43} - \theta\alpha_{41} - 1 & 0 \end{pmatrix}$$

and, with a particular choice of normalization,

$$\Phi_1^1 = \begin{pmatrix} \alpha_{11} & 0 & 0 & -\theta\alpha_{11} \\ \alpha_{21} - \alpha_{31} & 0 & 0 & -\theta(\alpha_{21} - \alpha_{31}) \\ \alpha_{31} - \alpha_{41} & 0 & 0 & -\theta(\alpha_{31} - \alpha_{41}) \\ \alpha_{41} & 0 & 0 & -\theta\alpha_{41} \end{pmatrix} = \begin{pmatrix} \alpha_{11} \\ \alpha_{21} - \alpha_{31} \\ \alpha_{31} - \alpha_{41} \\ \alpha_{41} \end{pmatrix} (1 \ 0 \ 0 \ -\theta).$$

4 Testing for common cyclical features in mixed-frequency VARs

In the two previous sections we have established some results about VECM representations of mixed-frequency VARs. This is important because the most popular common cycle approach (Vahid and Engle, 1993) aims at testing for the presence of common dynamics in VAR/VECM multivariate systems. The previous representations become the starting point that allows us to test for the presence of common cyclical features when two or more series are observed at different frequencies.

This being said, let us first point out that this paper focuses on the particular mixed-frequency case when m is relatively small, say 3 or 12. Indeed, we propose a method based on canonical correlations as well as Full Information Maximum Likelihood (FIML) techniques. It is well known that the latter has shortcomings when the number of variables is large compared to the number of observations T .

4.1 Definitions

Let us start by defining Ω as the set of all outcomes for the stochastic processes under consideration. Then, let Ω_t be the information set generated by the collection of sigma-fields $F_t = \sigma(Z_s, s \leq t)$, $t \geq 0$, and let Z_t be adapted to that filtration. In other words, Ω_t contains all the information present at moment t , implying that y_t and $x_t^{(m)}$ would be given, but $x_{t+1/m}^{(m)}$ would not.

We define the presence of common cycles in the mixed-frequency VAR (5) as follows:

Definition 6 *In a mixed-frequency VAR for the $m+1$ series ΔZ_t^* ,*

$$\Delta Z_t^* = \Phi_1 \Delta Z_{t-1}^* + \Phi_2 \Delta Z_{t-2}^* + \dots + \Phi_p^0 \tilde{Z}_{t-p} + \Phi_p^1 \tilde{Z}_{t-p} + \tilde{\varepsilon}_t,$$

with $\Delta Z_t^* = (\Delta y_t, \Delta_m x_t^{(m)}, \Delta_m x_{t-1/m}^{(m)}, \dots, \Delta_m x_{t-(m-1)/m}^{(m)})'$, a *Mixed-Frequency Serial Correlation Common Features (MF-SCCF)* arises iff there exists a matrix δ spanning the common feature space such that

$$\mathbb{E}(\delta' \Delta Z_t^* | \Omega_{t-1}) = 0.$$

This implies that there exists such a δ with the conditions $\delta' \Phi_1 = \dots = \delta' \Phi_{p-1} = \delta' \Phi_p^0 = \delta' \Phi_p^1 = \mathbf{0}$, i.e., it is a matrix that annihilates the short-run dynamics in the systems. In other words, δ lies in the intersection of the left null spaces of all Φ -matrices.

In particular, when δ is a $(m+1)$ -dimensional column vector normalized on the first element, we say that two series Δy_t and $\Delta_m x_t^{(m)}$ have a MF-SCCF relationship iff there exists a high-frequency-lag polynomial $\gamma(L_m)$ such that

$$\mathbb{E}([\Delta y_t - \gamma(L_m) \Delta_m x_t^{(m)}] | \Omega_{t-1}) = 0.$$

Hence, two series Δy_t and $\Delta_m x_t^{(m)}$ having a MF-SCCF imply that

$$\mathbb{E}([\Delta y_t - \gamma_1 \Delta_m x_t^{(m)}] | \Omega_{t-1}) = \gamma_2 \Delta_m x_{t-1/m}^{(m)} + \dots + \gamma_m \Delta_m x_{t-(m-1)/m}^{(m)}$$

and hence that there exists a (restricted) polynomial serial correlation common feature relationship of order m , *PSCCF*(m), as defined in Cubadda and Hecq (2001) between the two series.

4.2 Testing Procedure

Now that we have defined MF-SCCF, we can formulate the steps we use to test for mixed-frequency common cycles. Note that as far as testing for cointegration and testing for common cyclical features are concerned we follow a two-step procedure (Vahid and Engle, 1993, Hecq, Palm and Urbain, 2006), the validity of which has been established in Paruolo (2006). The steps are as follows:

1. We determine the lag length order in the level (or the log level) of the series for the mixed-frequency VAR in (1) to obtain p . We rely on usual multivariate information criteria (hereafter MIC), in this article we take AIC, SBC and HQ. Note that the performance of these criteria is investigated in our Monte Carlo section.
2. Given p , we test for the cointegrating rank $r = r_0 + r_1$. In doing so we rely on the Horvath and Watson (1995) approach because the $m-1 = r_0$ cointegrating vectors are prespecified.
3. In the absence of additional r_1 cointegrating vectors we perform a reduced rank regression in the transformed VAR(p) using a test on the zero eigenvalues computed from the canonical-correlation analysis

$$CanCor \left\{ \Delta Z_t^*, \begin{pmatrix} \Delta Z_{t-1}^* \\ \vdots \\ \Delta Z_{t-p+1}^* \\ \tilde{Z}_{t-p}^0 \end{pmatrix} \right\}, \quad (6)$$

where \tilde{Z}_{t-p}^0 is equivalent to \tilde{Z}_{t-p} except for the first and last elements:

$$\tilde{Z}_{t-p}^0 = \begin{pmatrix} 0 \\ \Delta_m x_{t-p}^{(m)} \\ \vdots \\ \Delta_m x_{t-p-(m-2)/m}^{(m)} \\ 0 \end{pmatrix}.$$

Furthermore, $CanCor\{V_t, W_t\}$ denotes the computation of canonical correlations between the two sets of variables V_t and W_t . The likelihood ratio type test, denoted by ξ_{LR} , considers the null hypothesis that there exist at most s common feature vectors against the alternative that there are less than s , i.e., $H_0 : rank(\delta) \leq s$ versus $H_A : rank(\delta) > s$. The test is obtained as

$$\xi_{LR} = -T \sum_{i=1}^s \ln(1 - \hat{\lambda}_i), \quad s = 1, 2, \quad (7)$$

where $\hat{\lambda}_i$ is the i^{th} smallest squared canonical correlation, i.e., the i^{th} smallest eigenvalue, computed from (6) above, i.e., from

$$\hat{\Sigma}_{VV}^{-1} \hat{\Sigma}_{VW} \hat{\Sigma}_{WW}^{-1} \hat{\Sigma}_{WV}, \quad (8)$$

or similarly from the symmetric matrix $\hat{\Sigma}_{VV}^{-1/2} \hat{\Sigma}_{VW} \hat{\Sigma}_{WW}^{-1} \hat{\Sigma}_{WV} \hat{\Sigma}_{VV}^{-1/2}$, where $\hat{\Sigma}_{ij}$ are the empirical covariance matrices with i and j being equal to V and/or W . In the mixed-frequency bivariate case and without counting the constant terms, the unrestricted transformed VAR has $(m+1)^2(p-1) + (m+1-2)(m+1)$ parameters. Note that $s = 1, 2$ because our 4-dimensional transformed VAR already contains two known cointegrating relationships such that s is maximum 2.

4. In the presence of cointegration, eigenvalues and eigenvectors, are computed from

$$CanCor \left\{ \Delta Z_t^*, \begin{pmatrix} \Delta Z_{t-1}^* \\ \vdots \\ \Delta Z_{t-p+1}^* \\ \tilde{Z}_{t-p} \end{pmatrix} \right\}, \quad (9)$$

where,

$$\tilde{Z}_{t-p} = \begin{pmatrix} y_{t-p} \\ \Delta_m x_{t-p}^{(m)} \\ \vdots \\ \Delta_m x_{t-p-(m-2)/m}^{(m)} \\ x_{t-p-(m-1)/m}^{(m)} \end{pmatrix},$$

and we focus on the test for $s = 1$. Indeed, due to the additional cointegrating relationship between the low- and high-frequency variables, there is only room for at most one common feature vector in this bivariate example. In the mixed-frequency bivariate case and without counting the constant terms, the unrestricted transformed VAR has $(m + 1)^2(p - 1) + (m + 1 - 2 + 1)(m + 1)$ parameters.

5. Alternatively and numerically equivalent is the estimation of the pseudo-structural form by FIML under the restriction that $\text{rank}(\delta) = 2$.⁶ In the present context, this representation might be easier for counting the number of restrictions as well as for considering additional restrictions on the common feature vectors. Let us write the pseudo-structural model in the presence of both sets of cointegrating vectors (prespecified and not prespecified) as

$$\begin{aligned}
& \begin{pmatrix} I & -\gamma \\ 0 & I \end{pmatrix} \Delta Z_t^* = \begin{pmatrix} 0 \\ \Lambda_1 \end{pmatrix} \Delta Z_{t-1}^* + \dots + \begin{pmatrix} 0 \\ \Lambda_{p-1} \end{pmatrix} \Delta Z_{t-(p-1)}^* + \begin{pmatrix} 0 \\ \Lambda_p \end{pmatrix} C^* \tilde{Z}_{t-p} + e_t \\
\Leftrightarrow & \begin{pmatrix} I_s & -\gamma_{s \times (m+1-s)} \\ 0_{(m+1-s) \times s} & I_{(m+1-s)} \end{pmatrix} \begin{pmatrix} \Delta y_t \\ \Delta_m x_t^{(m)} \\ \Delta_m x_{t-1/m}^{(m)} \\ \vdots \\ \Delta_m x_{t-(m-1)/m}^{(m)} \end{pmatrix} \\
= & \begin{pmatrix} \mathbf{0} \\ s \times (m+1) \\ \Lambda_1 \\ (m+1-s) \times (m+1) \end{pmatrix} \begin{pmatrix} \Delta y_{t-1} \\ \Delta_m x_{t-1}^{(m)} \\ \Delta_m x_{t-1-1/m}^{(m)} \\ \vdots \\ \Delta_m x_{t-1-(m-1)/m}^{(m)} \end{pmatrix} + \dots + \begin{pmatrix} \mathbf{0} \\ s \times (m+1) \\ \Lambda_{p-1} \\ (m+1-s) \times (m+1) \end{pmatrix} \begin{pmatrix} \Delta y_{t-p+1} \\ \Delta_m x_{t-(p-1)}^{(m)} \\ \Delta_m x_{t-(p-1)-1/m}^{(m)} \\ \vdots \\ \Delta_m x_{t-(p-1)-(m-1)/m}^{(m)} \end{pmatrix} \\
+ & \begin{pmatrix} \mathbf{0} \\ s \times (r_0+r_1) \\ \Lambda_p \\ (m+1-s) \times (r_0+r_1) \end{pmatrix} \underbrace{\begin{pmatrix} 1 & 0 & \dots & \dots & 0 & -\hat{\theta} \\ 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & \dots & \dots & 1 & 0 \end{pmatrix}}_{C^*} \begin{pmatrix} y_{t-p} \\ \Delta_m x_{t-p}^{(m)} \\ \Delta_m x_{t-p-1/m}^{(m)} \\ \vdots \\ \Delta_m x_{t-p-(m-2)/m}^{(m)} \\ x_{t-p-(m-1)/m}^{(m)} \end{pmatrix} + \begin{pmatrix} e_{1t} \\ e_{2t} \\ \vdots \\ e_{m+1t} \end{pmatrix},
\end{aligned}$$

where $\Lambda_1, \dots, \Lambda_p$ are the $(m + 1 - s)$ lower rectangular parts of the corresponding Φ -matrices that are going to be estimated by FIML. Note that the first row of C^* disappears in case $r_1 = 0$. This shows that we have to estimate $(m + 1 - s) \times (m + 1)$ parameters

⁶This can be easily achieved in Eviews or in OxMetrics, for instance.

in each of the Λ_i , $i = 1 \dots p - 1$. There are, in the presence of one additional cointegrating vector, $(m + 1 - s) \times m$ parameters in Λ_p . Finally, there are $s \times (m + 1 - s)$ parameters in γ . Without additional cointegrating vectors, Λ_p has only $(m + 1 - s) \times (m - 1)$ parameters.

6. The null and alternative hypotheses of the common feature test in terms of the pseudo-structural form become

$$\begin{aligned} H_0 &: \delta' \Phi_1 = \dots = \delta' \Phi_p = 0 \\ H_A &: \delta' \Phi_1 = \dots = \delta' \Phi_p \neq 0, \end{aligned}$$

where the rank of δ is implicitly equal to either 1 or 2 as argued before. The test is simply performed using the following likelihood ratio type test:

$$LR_s = 2(l_{H_A} - l_{H_0}) \sim \chi^2(v),$$

where l_{H_A} and l_{H_0} denote the value of the log-likelihood function evaluated under the alternative and the null hypothesis, respectively. The degrees of freedom v is the difference between the number of parameters in the unrestricted mixed-frequency VAR (with cointegration) and the number of coefficients in the pseudo-structural system.

7. Finally, additional restrictions on the vectors of γ can be imposed if necessary for extracting a meaningful relationship between y and x .

The next section helps to digest all these steps.

5 Applications

To illustrate the methods discussed in the previous sections, we extract two pairs of time series, once for the U.S. and once for Germany.⁷ In both cases y is the quarterly real gross national product (ref. GNPC96 and B1GM for the U.S. and Germany, respectively), seasonally adjusted. For the U.S. the series is observed from 1948Q1 until 2011Q4 whereas for Germany it is observed from 1991Q1 until 2012Q1. For x we consider, for the same time span, the monthly seasonally adjusted industrial production index (ref. INDPRO and PROD, respectively), or IPI from here on. Since GNP is a quarterly variable and IPI is observed on a monthly basis we have that $m = 3$.

5.1 Common Feature Analysis

As a first step, we determine the VAR order in the log-level of the series by relying on MIC. That implies that we compute the mixed-frequency VAR à la Ghysels (2012) disregarding the possibility that the series may be I(1) and possibly cointegrated. After assuming the I(1)-ness of both series involved (we indeed do not reject the null of a unit root for all series under consideration), we may test for the cointegrating rank r . As discussed above, due to the presence

⁷From respectively <http://alfred.stlouisfed.org/> and <http://ec.europa.eu/eurostat/>.

of r_0 prespecified cointegrating vectors, we follow the approach of Horvath and Watson (1995). In particular, since $m = 3$ we test the null hypothesis that merely the two known cointegrating vectors are present against the alternative of one additional cointegrating vector to capture the long-run relationship between GNP and IPI (Götz et al., 2012a). We can formulate the likelihood ratio test statistic as

$$LR_{r_0, r_1} = 2(l_{H_A} - l_{H_0}) = 2(l_{r=r_0+r_1} - l_{r=r_0}),$$

where l_{H_0} and l_{H_A} denote the value of the log-likelihood function evaluated under the null and the alternative hypothesis, respectively, i.e., when once only the prespecified cointegrating vectors are imposed and once the prespecified ones and one capturing potential cointegration between y and x . In terms of our empirical applications, this approach involves the log-likelihood of the following two regressions:

$$\text{Under } H_0 : \begin{pmatrix} \Delta GNP_t \\ \Delta IPI_t^{(3)} \\ \Delta IPI_{t-1/3}^{(3)} \\ \Delta IPI_{t-2/3}^{(3)} \end{pmatrix} \text{ on } \underbrace{\begin{pmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}}_{\begin{pmatrix} \Delta_m IPI_{t-1}^{(3)} \\ \Delta_m IPI_{t-1-1/3}^{(3)} \end{pmatrix}} \begin{pmatrix} GNP_{t-1} \\ IPI_{t-1}^{(3)} \\ IPI_{t-1-1/3}^{(3)} \\ IPI_{t-1-2/3}^{(3)} \end{pmatrix}, \begin{pmatrix} \Delta GNP_{t-1} \\ \Delta IPI_{t-1}^{(3)} \\ \Delta IPI_{t-1-1/3}^{(3)} \\ \Delta IPI_{t-1-2/3}^{(3)} \end{pmatrix}$$

and

$$\text{Under } H_A : \begin{pmatrix} \Delta GNP_t \\ \Delta IPI_t^{(3)} \\ \Delta IPI_{t-1/3}^{(3)} \\ \Delta IPI_{t-2/3}^{(3)} \end{pmatrix} \text{ on } \underbrace{\begin{pmatrix} 1 & -\theta & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}}_{\begin{pmatrix} GNP_{t-1} - \theta IPI_{t-1}^{(3)} \\ \Delta_m IPI_{t-1}^{(3)} \\ \Delta_m IPI_{t-1-1/3}^{(3)} \end{pmatrix}} \begin{pmatrix} GNP_{t-1} \\ IPI_{t-1}^{(3)} \\ IPI_{t-1-1/3}^{(3)} \\ IPI_{t-1-2/3}^{(3)} \end{pmatrix}, \begin{pmatrix} \Delta GNP_{t-1} \\ \Delta IPI_{t-1}^{(3)} \\ \Delta IPI_{t-1-1/3}^{(3)} \\ \Delta IPI_{t-1-2/3}^{(3)} \end{pmatrix}$$

According to the Schwarz and the Hannan-Quinn (SC and HQ, respectively) information criterion a lag length of 1 is chosen for both countries under consideration. A VAR(2) is preferred for the U.S. using the Akaike information criterion (AIC). Consequently, we carry out the analysis with both lag lengths for the US in order to evaluate the robustness of the results.

It turns out that $LR_{r_0, r_1}^{US} = 8.62$ in the VAR(2) and 8.88 for the VAR(1) whereas $LR_{r_0, r_1}^{GER} = 25.41$ in the VAR(1) case such that with a Horvath-Watson (1995)-critical value of 18.37 we do not reject the null hypothesis of no cointegration between GNP and IPI for the US, but do reject it for Germany. The cointegrating vector from the restricted VECM in the Johansen's

approach corresponding to r_1 is

$$\text{Germany} : GNP_t - \underset{(0.048)}{0.751} IPI_{t-2/3}^{(3)} \sim I(0)$$

This result will have an impact on how to set up the transformed VAR for both applications which, in turn, has an implication on how to test for common cyclical features thereafter (see examples 4 and 5 from Section 3).

Let us now look at the presence of additional common cyclical feature restrictions. For the US, we first look at the VAR(1) in the log-level of the series which corresponds to the following transformed mixed-frequency VAR:

$$\begin{pmatrix} \Delta GNP_t \\ \Delta_m IPI_t^{(3)} \\ \Delta_m IPI_{t-1/3}^{(3)} \\ \Delta_m IPI_{t-2/3}^{(3)} \end{pmatrix} = \begin{pmatrix} 0.005 \\ 0.001 \\ 0.002 \\ 0.001 \end{pmatrix} + \begin{pmatrix} 0 & 0.488 & 0.265 & 0 \\ 0 & 0.247 & 0.023 & 0 \\ 0 & 0.281 & 0.081 & 0 \\ 0 & 0.362 & 0.131 & 0 \end{pmatrix} \begin{pmatrix} GNP_{t-1} \\ \Delta_m IPI_{t-1}^{(3)} \\ \Delta_m IPI_{t-1-1/3}^{(3)} \\ IPI_{t-1-2/3}^{(3)} \end{pmatrix} + e_t$$

with a value of the log-likelihood of 3477.474. Because the rank of the transformed VAR coefficient matrix is 2, we know that, by construction, its orthogonal complement and then the common feature space δ is also of rank 2 in this four-dimensional system. Hence, the normalized common feature matrix that takes the form

$$\delta' = \begin{pmatrix} 1 & 0 & \times & \times \\ 0 & 1 & \times & \times \end{pmatrix},$$

where \times denotes a parameter to be estimated, always exists. Premultiplying the VAR by

$$\begin{pmatrix} \delta' \\ 0_{(2 \times 2)} & I_{(2)} \end{pmatrix}$$

leads to the pseudo-structural form with the same likelihood because no restrictions are added and hence no test is possible.

We may further restrict δ and investigate whether there exists a common cyclical feature relationship between the growth rates of GNP and IPI. Restricting for instance the first possible relationship as in

$$\delta' = \begin{pmatrix} 1 & 0 & 0 & \times \\ 0 & 1 & \times & \times \end{pmatrix}$$

yields a log-likelihood of 3477.067 compared to 3477.474 from the previous mixed-frequency VAR imposing no common feature restrictions. This obviously means a non-rejection of the additional null hypothesis, $H_0 : \delta' \Phi_1 = \dots, \delta' \Phi_p = 0$ (see section 5.1). The MF-SCCF relationship is now given by:

$$\Delta GNP_t = \underset{(0.0001)}{0.005} + \underset{(0.125)}{1.509} \Delta_m IPI_{t-2/3}^{(3)} + w_n,$$

i.e., a relationship between the quarterly growth rate of GNP and the monthly difference of IPI for the first month of the corresponding quarter. Still for the US, but with a VAR(2), as favoured by the AIC, we have a log-likelihood of 3480.36 in

$$\begin{pmatrix} \Delta GNP_t \\ \Delta_m IPI_t^{(3)} \\ \Delta_m IPI_{t-1/3}^{(3)} \\ \Delta_m IPI_{t-2/3}^{(3)} \end{pmatrix} = \begin{pmatrix} 0.005 \\ 0.001 \\ 0.001 \\ 0.000 \end{pmatrix} + \begin{pmatrix} 0.043 & 0.478 & 0.255 & 0.004 \\ 0.353 & 0.175 & -0.010 & -0.311 \\ 0.259 & 0.227 & 0.045 & -0.031 \\ 0.113 & 0.330 & 0.100 & 0.021 \end{pmatrix} \begin{pmatrix} \Delta GNP_{t-1} \\ \Delta_m IPI_{t-1}^{(3)} \\ \Delta_m IPI_{t-1-1/3}^{(3)} \\ \Delta_m IPI_{t-1-2/3}^{(3)} \end{pmatrix} \\ + \begin{pmatrix} 0 & -0.016 & 0.027 & 0 \\ 0 & -0.085 & 0.011 & 0 \\ 0 & -0.096 & -0.046 & 0 \\ 0 & 0.017 & -0.035 & 0 \end{pmatrix} \begin{pmatrix} GNP_{t-2} \\ \Delta_m IPI_{t-2}^{(3)} \\ \Delta_m IPI_{t-2-1/3}^{(3)} \\ IPI_{t-2-2/3}^{(3)} \end{pmatrix} + e_t.$$

In this case there does not necessarily exist a matrix that annihilates all of the dynamics, namely the two matrices. For $s = 1$ we have a log-likelihood value of 3479.877, for $s = 2$ we have a log-likelihood value of 3474.863 in the pseudo-structural systems estimated by FIML. Consequently, we do not reject the null of $s = 2$ common feature vectors with a likelihood ratio test of $2(3480.36 - 3474.86) = 11 \sim \chi^2(8)$. In addition to $s = 2$, we are also able to restrict the first common feature vector as in the ($p = 1$)-case; The log-likelihood with this additional restriction is 3474.732. The relationship is

$$\Delta GNP_t = \underset{(0.0001)}{0.005} + \underset{(0.128)}{1.488} \Delta_m IPI_{t-2/3}^{(3)} + v_n,$$

such that very similar results as before emerge, but now with a formal test for MF-SCCF.

For Germany with $r_0 + r_1$ cointegrating vectors we have:

$$\begin{pmatrix} \Delta GNP_t \\ \Delta_m IPI_t^{(3)} \\ \Delta_m IPI_{t-1/3}^{(3)} \\ \Delta_m IPI_{t-2/3}^{(3)} \end{pmatrix} = \begin{pmatrix} 0.102 \\ -1.594 \\ -1.008 \\ 0.166 \end{pmatrix} + \begin{pmatrix} 0 & 0.297 & 0.147 & 0 \\ 0 & 0.167 & -0.093 & 0 \\ 0 & 0.143 & 0.067 & 0 \\ 0 & 0.043 & 0.137 & 0 \end{pmatrix} \begin{pmatrix} GNP_{t-1} \\ \Delta_m IPI_{t-1}^{(3)} \\ \Delta_m IPI_{t-1-1/3}^{(3)} \\ IPI_{t-1-2/3}^{(3)} \end{pmatrix} \\ + \begin{pmatrix} -0.010 \\ 0.162 \\ 0.103 \\ -0.017 \end{pmatrix} (GNP_{t-1} - 0.751 IPI_{t-1-2/3}^{(3)}) + u_t.$$

Due to the ranks of these matrices, we know that there always exists a vector that annihilates the dynamics. This is an application of the mixed form reduced rank approach of Hecq et al. (2006). The common feature relationship and the standard errors (in brackets beneath the estimators) obtained by FIML, however, show that there are no cofeature relationships between

the GNP and IPI series alone:

$$\Delta GNP_t = 0.006 + \underset{(0.030)}{2.596} \Delta_m IPI_t^{(3)} - \underset{(24.381)}{3.411} \Delta_m IPI_{t-1/3}^{(3)} + \underset{(27.25)}{4.912} \Delta_m IPI_{t-2/3}^{(3)} + w_n,$$

such that the only common cyclical feature relationship involves the growth rate of the gross national product which is consequently a white noise process.

Alternative to a common feature analysis in a mixed-frequency setup, the high-frequency variables can be temporally aggregated in order to work with a common (low-) frequency VAR. Note, however, that common features are not invariant to temporal aggregation. Time aggregation of high-frequency variables may not only cause common feature relationships, that are present in mixed frequencies, to disappear in a common frequency framework, it may also create additional cofeature vectors (Marcellino, 1999).

Two temporal aggregation schemes are considered here, point-in-time and average sampling. The former takes a specific high-frequency observation as the corresponding low-frequency one and is usually applied to stock variables. The latter simply uses the average of the high-frequency observations as the corresponding t -observation and is usually of interest when flow variables are dealt with (see Marcellino, 1999, or Silvestrini and Veredas, 2008). We only consider the U.S. data for the common feature analysis in these low-frequency VARs. The resulting elasticities (standard errors in brackets) of ΔGNP with respect to $\Delta IPI_t^{(3)}$, $\Delta IPI_{t-1/3}^{(3)}$, $\Delta IPI_{t-2/3}^{(3)}$ and $\overline{\Delta IPI_t^{(3)}}$ are 0.7187 (0.0696), 0.4033 (0.0364), 0.279 (0.016) and 0.4097 (0.0328), respectively, where $\overline{IPI_t^{(3)}}$ refers to the average of the monthly IPI-observations in period t . The corresponding common feature test statistics computed from the pseudo-structural forms of the low-frequency VARs are 10.09, 7.9, 31 and 4.48 for the respective temporal aggregation approaches. At the 5% significance level, the null hypothesis of one common feature is rejected in all cases, implying that there exist two independent, i.e., non-common and non-synchronous, cycles for two variables under consideration. In other words, the commonalities that are present in the mixed-frequency setup are not visible anymore after time aggregation. Finally, note that the cofeature elasticities are also very different to the mixed-frequency setup.

5.2 Forecasting

In the remainder of this section a small forecasting exercise is presented to emphasize the potential benefits of employing a mixed-frequency approach when dealing with common cycles. Potential gains of incorporating common feature restrictions for forecasting already exist in the literature. Vahid and Issler (2002) show that the costs in terms of forecast accuracy associated with ignoring common feature relationships in VARs may be considerable. Athanasopoulos, Guillén, Issler and Vahid (2011) also show that models imposing restrictions associated with common cycles outperform unrestricted VAR models for different sample sizes, especially if the data are generated by comparably severe rank restrictions. Anderson and Vahid (2010) argue that VARs "with cointegration and common cycles can be [...] viewed as observable factor models [...that...] have potential for forecasting" (page 5). In a set of Monte Carlo

studies the authors illustrate that over- and underspecification of the associated rank restrictions deteriorate the forecast performance in almost all cases. While the cointegration structure seems to be more crucial in general, mis-specification of the common cycle rank restrictions particularly worsens the forecast accuracy in the short-run. In their conclusion, Anderson and Vahid (2010) advise to consider reduced rank error-correction models for forecasting if there are compelling reasons to consider common trends or cycles in a particular data set.

We consider 28 out-of-sample one-step-ahead forecasts of the real U.S. GNP growth rate corresponding to the period from 2005Q1 until 2011Q4. We recursively update the estimation period and select the lag length at each moment in time. Furthermore, we test for the number of common feature vectors at each moment in time. Note that we do not use real-time data, but final-vintage data. Indeed, real-time data would require a vintage-specific analysis of which models to employ, which lag length to choose, cointegration properties and so on (see Götz et al, 2012b for an example of mixed-frequency models in a real-time data set). As already investigated above, the null hypothesis of no cointegration cannot be rejected for the U.S. such that we assume no cointegration to be present for each estimation period under consideration.

We are interested in the forecast accuracy of three mixed-frequency stationary VAR models. First, the model without restrictions on the short-run parameters, denoted by MF_VAR1_FULL. Second, the model imposing zero-restrictions corresponding to the absence of cointegration between GNP and IPI in the transformed VAR, denoted by MF_VAR1_ZEROS. Finally, the model imposing, on top of the just mentioned zero-restrictions, a set of common feature vectors, denoted by MF_VAR1_CF. In order to investigate whether the development of such mixed-frequency VAR models leads to an improved forecast performance compared to VARs stemming from traditional aggregation methods, we confront our three MF-VARs with two low-frequency VARs, resulting from applying average sampling to IPI on the one hand and point-in-time sampling using the end-of-period observations, $IPI_t^{(3)}$, on the other hand. In both cases a VAR in first differences, denoted by AV_VAR1_DIFF and PIT_VAR1_DIFF, respectively, are considered. RMSFE errors are computed for each model.

As far as the lag length selection is concerned, it turns out that for both, the mixed-frequency and time aggregated models, a VAR(1) model in first differences is chosen by the Schwarz or Hannan-Quinn information criteria at each moment in time. When testing for common features in the mixed-frequency VARs, the null hypothesis of $s = 2$ common feature vectors could not be rejected for all estimation periods considered. Hence, MF_VAR1_CF imposes two common feature vectors at each step of the forecasting exercise, whereby the first vector is further restricted as done in the previous subsection. Even though no common feature could have been detected at the 5% significance level at each moment in time, we include the models imposing one common feature vector into the analysis, i.e., AV_VAR1_1CF and PIT_VAR1_1CF. As a justification, although no common feature was detected for average sampled IPI at the 5% significance level, the null hypothesis of one common feature could not be rejected at the 1% significance level.

Figure 1 shows the forecasts of the real U.S. GNP growth rate from 2005Q1 to 2011Q4 corresponding to each of the seven models just introduced as well as the actual data (solid

black line). Note that the black dotted or dashed lines correspond to the mixed-frequency transformed VARs, the grey dotted or dashed ones to the VARs based on averaging the high-frequency variable and, finally, the grey solid lines to the VARs based on taking the last monthly observation of IPI per quarter (point-in-time).



Figure 1: Out-of-sample one-step-ahead static forecasts of real GNP growth in the U.S. from 2005Q1 to 2011Q4

It becomes clear from Figure 1 that the forecasts based on the mixed-frequency models are, compared to the competing time aggregated models, able to track the actual data rather well, especially the trough associated with the financial crisis in 2008. The forecasts based on VARs stemming from temporally aggregating the high-frequency variables severely underestimate this drop in the GNP growth rate. This is reflected in the RMSFEs summarized in Table 1.

Table 1: RMSFEs of the seven competing models

Model	RMSFE
MF_VAR1_ZEROS	0.005671
MF_VAR1_CF	0.005672
MF_VAR1_FULL	0.005766
PIT_VAR1_DIFF	0.006527
PIT_VAR1_1CF	0.006591
AV_VAR1_1CF	0.007589
AV_VAR1_DIFF	0.00766

Note: The figures represent the RMSFEs of the seven competing models: The model 1) involving no restrictions, 2) imposing zero-restrictions, 3) imposing zero-restrictions and two common features, 4)-5) based on point-in-time sampling with and without a common feature imposed, respectively, and 5)-6) based on averaging with and without a common feature imposed.

As becomes evident from the figures, the forecast accuracy of the mixed-frequency transformed VARs is significantly better than the one of the VARs based on classical temporal aggregation methods. Hence, in the case of predicting the real U.S. GNP growth rate from 2005Q1 to 2011Q1 it seems advantageous to consider mixed-frequency VAR models instead of temporally aggregating the high-frequency variable.

6 Monte Carlo study

Taking the preceding small forecasting exercise as a warm-up, we want to investigate the behavior of mixed-frequency VAR models more closely in this section. Naturally, we are interested in the common feature test behavior, i.e., in the size and power of the LR type test in (7) or the one corresponding to (10), depending on whether there are additional cointegrating vectors above the prespecified ones. Since the computation of the test statistic, however, is dependent on the lag length p determined by MIC (see section 4), we also analyze the frequencies with which these criteria select different lag lengths.

To this end we conduct a Monte Carlo study, in which we simulate a mixed-frequency VAR that matches some of the characteristics that we have obtained using U.S. data, namely a VAR of order $p = 2$ in the level of the I(1) series and no additional cointegrating vectors above the known ones. Hence, it is a two-variable system with $m = 3$ for the high-frequency series, in

which the following parameter values in the transformed VAR representation are chosen:

$$\begin{pmatrix} \Delta y_t \\ \Delta_m x_t^{(3)} \\ \Delta_m x_{t-1/3}^{(3)} \\ \Delta_m x_{t-2/3}^{(3)} \end{pmatrix} = \begin{pmatrix} 0.05 \\ 0.05 \\ 0.05 \\ 0.05 \end{pmatrix} + \begin{pmatrix} 0.2 & 0.5 \\ -0.2 & 0.2 \\ 0.4 & 0.5 \\ 0.5 & 0.25 \end{pmatrix} \begin{pmatrix} 0.5 & 0.1 \\ 0.6 & 0.5 \\ 0.2 & -0.5 \\ 0.1 & 0.25 \end{pmatrix}' \begin{pmatrix} \Delta y_{t-1} \\ \Delta_m x_{t-1}^{(3)} \\ \Delta_m x_{t-1-1/3}^{(3)} \\ \Delta_m x_{t-1-2/3}^{(3)} \end{pmatrix} \\ + \begin{pmatrix} 0 & 0.2 & 0.5 & 0 \\ 0 & -0.2 & 0.2 & 0 \\ 0 & 0.4 & 0.5 & 0 \\ 0 & 0.5 & 0.25 & 0 \end{pmatrix} \begin{pmatrix} y_{t-2} \\ \Delta_m x_{t-2}^{(3)} \\ \Delta_m x_{t-2-1/3}^{(3)} \\ x_{t-2-2/3}^{(3)} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \varepsilon_{4t} \end{pmatrix},$$

where

$$\varepsilon_t \sim \text{i.i.d. } N(\mathbf{0}, \Sigma) \text{ with } \Sigma = \begin{pmatrix} 1 & 0.5 & 0.5 & 0.5 \\ 0.5 & 1 & 0.9 & 0.9 \\ 0.5 & 0.9 & 1 & 0.9 \\ 0.5 & 0.9 & 0.9 & 1 \end{pmatrix}.$$

The normalized cofeature matrix that annihilates the two coefficient matrices of this DGP are fixed to:

$$\delta' = \begin{pmatrix} 1 & 0 & -\frac{4}{3} & \frac{2}{3} \\ 0 & 1 & -1 & 1.2 \end{pmatrix}.$$

Note that because the DGP is generated as mixed-frequency transformed VAR, we know the correct lag length to be $p = 2$ and the number of common feature vectors to be $s = 2$. Consequently, we will investigate the extent to which multivariate information criteria choose the correct lag length instead of under- or overestimating it. Likewise, given a lag length p , we can analyze size ($s = 2$) and power ($s = 3$ or $s = 4$) of the LR type test in (7).

As already mentioned before, an alternative to considering a mixed-frequency VAR is temporally aggregating the high-frequency variables in order to work in a common (low-)frequency setup. Consequently, we mimic the scenario in which practitioners are unaware of the mixed-frequency nature of the underlying data and work with time aggregated IPI data. Assuming they follow the same testing procedure, i.e., relying on MIC to choose p and using (7) to test for common feature relationships, employing their low-frequency system, it is of interest to analyze in how far the use of time aggregated data instead of high-frequency ones changes the results.

Note that as soon as data are temporally aggregated, the dynamics of the underlying system change completely (see Marcellino, 1999, Silvestrini and Veredas, 2008, or Lütkepohl, 2005). Indeed, an n -dimensional VAR(p) can be written in a final equation system (Zellner and Palm, 1974) such that each component is at most an ARMA($np, (n-1)p$) process. In the presence of s SCCF-restrictions the n -dimensional VAR(p) yields at most ARMA($(n-s)p, (n-s)p$) models (Cubadda, Hecq and Palm, 2009). In our case, the VAR(2) becomes a VARMA(4, 4) after temporal aggregation of IPI (because $n = 4$ and $s = 2$), that can be written as VAR(∞). For practical reasons, the VARMA(4, 4) is usually approximated by a finite order VAR. Hence,

because the true lag length is infinity in this case, we can merely report the frequencies with which MIC detect the different lag lengths under consideration. Note that we do not derive the final equation system corresponding to the mixed-frequency transformed VAR because it depends not only on the particular parameter values chosen, but also on T , making it a tedious and challenging task that could be tackled in future work.

As already mentioned before, time aggregation also has an impact on the analysis of common cycles (Marcellino, 1999). For example, it may cause common feature relationships, that are present in mixed frequencies, to disappear in a common frequency framework as illustrated at the end of section 5.1. In our case, temporal aggregation of IPI leads to a 2-dimensional low-frequency VAR, for which the corresponding common feature test can be performed and the rejection frequencies for $s = 1$ or $s = 2$ can be computed. The rejection frequencies for $s = 2$ represent power against a bivariate white noise system. The ones for $s = 1$, however, cannot clearly be interpreted as size or power, because we do not know whether one synchronous common feature is present after time aggregation (in which case it would measure size) or none (in which case it would correspond to power). Again, the precise impact of temporal aggregation in a mixed-frequency VAR on common cycles analyses is left for further research.

For both, the mixed-frequency and time aggregated models, we consider $T = 40$ and 100 corresponding to respectively 10 and 25 years of quarterly data. An additional burning period of 100 observations is used to initialize the process. All figures are based on 10,000 replications. We concentrate on: (i) The frequencies with which different lag lengths are detected by MIC and (ii) the rejection frequencies for common feature test statistics corresponding to the likelihood ratio type tests in (7).

For the mixed-frequency models, Table 2 reports the frequencies with which MIC (AIC, SC and HQ) select different lag lengths. We estimate mixed-frequency VARs of orders from one to six but without imposing the zero-restrictions in the last matrix, i.e. in Φ_p for $p = 1, \dots, 6$. These models are denoted by MF_VAR. Models imposing zero-restrictions are denoted by MF_VAR_0. It is of interest to investigate whether non-imposition of the zero-restrictions in MF_VAR causes MIC to underestimate the lag length, i.e., choose $p = 1$ instead of $p = 2$.

Table 2: Frequencies with which MIC select different lag lengths; Mixed-frequency models

		$T = 40$						
		p	1	2	3	4	5	6
MF_VAR	AIC		0	43.53	4.04	2.87	5.24	44.32
	SC		0.04	90.11	2.32	0.76	0.89	5.88
	HQ		2.5	97.5	0	0	0	0
MF_VAR_0	AIC		0	63.4	5.4	1.8	4.4	25
	SC		0	95.9	1.8	0.3	0.2	1.8
	HQ		0.4	99.6	0	0	0	0
		$T = 100$						
MF_VAR	AIC		0	96.2	2.78	0.52	0.2	0.3
	SC		0	100	0	0	0	0
	HQ		0	100	0	0	0	0
MF_VAR_0	AIC		0	95.6	3.6	0.5	0.1	0.2
	SC		0	100	0	0	0	0
	HQ		0	100	0	0	0	0

Note: The figures represent frequencies with which the Akaike (AIC), Schwarz (SC) and Hannan-Quinn (HQ) information criterion detect different lag lengths. The DGP is a mixed-frequency transformed VAR(2) in the level of two I(1)-series, where no additional cointegrating relationships above the known ones and two common feature vectors are present. Models with VAR orders from 1 to 6 are estimated where zero-restrictions, corresponding to the absence of additional cointegrating relationships, are once imposed (denoted MF_VAR_0) and once not imposed (denoted MF_VAR). The sample size is $T = 40$ (top) or $T = 100$ (bottom).

We observe from Table 2 that the correct lag length, i.e., $p = 2$, is accurately chosen in the mixed-frequency transformed VARs, especially using the HQ and to a slightly lesser extent by SC. AIC overestimates the dynamics for $T = 40$. For all information criteria, imposing the zero-restrictions increases the frequencies to detect the correct lag length. Note, however, that the model not imposing the zero-restrictions does not underestimate p . Furthermore, the frequency to obtain the correct lag length quickly increases with the sample size.

Table 3 contains the corresponding results for the low-frequency models obtained by temporally aggregating the high-frequency observations. The two time aggregation methods already introduced before are considered, i.e., average and point-in-time sampling. Note that as far as the latter is concerned, we only report the results for point-in-time sampling on the end-of-period (PiT $x_t^{(3)}$) or start-of-period (PiT $x_{t-2/3}^{(3)}$) observations to save space (results available upon request).

Table 3: Frequencies with which MIC select different lag lengths; Low-frequency models

		$T = 40$						
		p	1	2	3	4	5	6
Average	AIC		19.2	51.18	12.39	7.81	4.31	5.11
	SC		35.48	53.6	6.56	2.82	0.87	0.67
	HQ		60.2	38.29	1.2	0.28	0.02	0.01
PiT $_x^{(3)}$	AIC		5.96	48.5	17.7	13.4	6.97	7.47
	SC		16.41	63.8	7.41	8.77	2.19	1.42
	HQ		40.2	56.4	2.13	1.14	0.08	0.05
PiT $_x^{(3)}$ $_{t-2/3}$	AIC		28.38	26.2	20	11.78	5.68	7.96
	SC		52.17	27.48	12.8	4.92	1.5	1.13
	HQ		79.78	16.5	3.02	0.62	0.07	0.01
		$T = 100$						
Average	AIC		0.58	62.3	18.59	13.18	3.45	1.9
	SC		4.37	85.9	7.35	2.22	0.12	0.04
	HQ		17.3	81.7	0.92	0.08	0	0
PiT $_x^{(3)}$	AIC		0	25.1	11.52	48.42	8.84	6.12
	SC		0.17	66.9	10.01	21.3	1.28	0.34
	HQ		2.79	92	3	2.18	0.03	0
PiT $_x^{(3)}$ $_{t-2/3}$	AIC		3.1	19.2	31.1	29.7	8.65	8.25
	SC		20.7	41.38	25.6	11.1	0.95	0.27
	HQ		55.6	36.8	6.76	0.83	0.01	0

Note: The figures represent frequencies with which the Akaike (AIC), Schwarz (SC) and Hannan-Quinn (HQ) information criterion detect different lag lengths. The DGP is a mixed-frequency transformed VAR(2) in the level of two I(1)-series, where no additional cointegrating relationships above the known ones and two common feature vectors are present. Common low-frequency models obtained via three different temporal aggregation methods are considered: (a) Average sampling (denoted Average), (b) point-in-time sampling using the end-of-period observation (denoted PiT $_x^{(3)}$) and (c) point-in-time sampling using the start-of-period observation (denoted PiT $_x^{(3)}$ $_{t-2/3}$). The sample size is $T = 40$ (top) or $T = 100$ (bottom).

It can be seen from Table 3 that, consistently with the previous discussion, depending on the temporal aggregation method, the MIC and the value of T under consideration, a different lag length is chosen. This is in contrast to the mixed-frequency models, for which we know the true value of p .

Concerning the common feature test statistics, we rely on the estimation of the squared canonical correlation reviewed in Section 4. For the mixed-frequency models, Table 4 reports the rejection frequencies associated with the likelihood ratio type test in (7). As argued before, the size of the test should be interpreted as the rejection frequency for $s = 2$, power can be

observed for $s = 3$ and $s = 4$.

Table 4: Rejection frequencies of the LR type test in (7); Mixed-frequency models

		$T = 40$						
		p	1	2	3	4	5	6
MF_VAR	$s = 1$		0.23	0.58	0.99	2.35	5.69	15.4
	$s = 2$		1.42	8.22	17.7	25.4	44.3	68.5
	$s = 3$		20.4	75.7	79.1	87.3	94.2	98.7
	$s = 4$		100	100	100	100	100	100
MF_VAR_0	$s = 1$		-	0.37	0.74	1.45	3.8	9.41
	$s = 2$		-	6.49	10.3	18.9	34.2	55.9
	$s = 3$		18	77.5	76.6	82.8	90.8	97.3
	$s = 4$		99.8	100	100	100	100	100
		$T = 100$						
MF_VAR	$s = 1$		0.24	0.26	0.29	0.43	0.53	0.79
	$s = 2$		1.98	6.59	8.17	9.87	12.4	16.6
	$s = 3$		37	99.4	98.5	97.4	96.9	96.9
	$s = 4$		100	100	100	100	100	100
MF_VAR_0	$s = 1$		-	0.32	0.26	0.41	0.53	0.73
	$s = 2$		-	5.68	7.55	9.22	11.1	14.6
	$s = 3$		45.6	99.7	99	97.9	97.1	96.8
	$s = 4$		100	100	100	100	100	100

Note: The figures represent rejection frequencies associated with the likelihood ratio type test in equation (7). The DGP is a VAR(2) in the level of two I(1)-series, where no additional cointegrating relationships above the known ones and two common feature vectors are present. Models with VAR orders from 1 to 6 are estimated where zero-restrictions, corresponding to the absence of additional cointegrating relationships, are once imposed (denoted MF_VAR_0) and once not imposed (denoted MF_VAR). The sample size is $T = 40$ (top) or $T = 100$ (bottom). Size should be interpreted as the rejection frequency for $s = 2$, power can be observed for $s = 3$ and $s = 4$.

Using the correct lag length $p = 2$, the size distortion is small, especially if one correctly accounts for the zeros in the last coefficient matrix. Size distortions are amplified when the lag length is overestimated. This is reflected by the figures in the columns corresponding to $p > 2$. As discussed above, however, we advise to rely on HQ or SC to get the correct lag length for the mixed-frequency models. The test is undersized when p is underestimated. With respect to the power of the test, note that for $T = 40$ we reject the presence of 3 common feature vectors 75% of the cases when the lag length is chosen correctly and up to almost 100% of the cases when it is overestimated. When the $p = 1$, however, the power deteriorates. The presence of 4 common feature vectors, however, is rejected in all cases. As soon as the sample size increases, we reject both $s = 3$ and $s = 4$ in almost all cases (except for $p = 1$ and $s = 3$).

Table 5 reports the rejection frequencies of the low-frequency counterpart to the common feature test statistic in (7). As argued before, note that the frequencies corresponding to $s = 2$ represent power against a bivariate white noise system, whereas the ones for $s = 1$ cannot clearly be interpreted as size or power.

Table 5: Rejection frequencies of the LR type test in (7); Low-frequency models

		$T = 40$						
		p	1	2	3	4	5	6
Average	$s = 1$		3.82	7.15	8.08	10.3	11.1	13.3
	$s = 2$		68.4	88.5	87.4	87.4	86.6	86.8
PiT $x_t^{(3)}$	$s = 1$		4.54	9.05	9.83	14.1	14.4	16.5
	$s = 2$		81.3	97.2	96.2	97.6	97.1	96.8
PiT $x_{t-2/3}^{(3)}$	$s = 1$		4.97	22.5	23.8	25.4	25.6	28.3
	$s = 2$		89.2	92.8	95.2	95	94.3	94.8
		$T = 100$						
Average	$s = 1$		4.43	8.8	10.7	12.2	12.1	12.4
	$s = 2$		97.7	100	100	100	100	100
PiT $x_t^{(3)}$	$s = 1$		11.9	21.3	22.1	30.8	28	27.5
	$s = 2$		100	100	100	100	100	100
PiT $x_{t-2/3}^{(3)}$	$s = 1$		5.79	55.7	58.5	56.3	54.2	53.3
	$s = 2$		100	100	100	100	100	100

Note: The figures represent rejection frequencies associated with the likelihood ratio type test in equation (7). The DGP is a VAR(2) in the level of two I(1)-series, where no additional cointegrating relationships above the known ones and two common feature vectors are present. Common low-frequency models obtained via three different temporal aggregation methods are considered: (a) Average sampling (denoted Average), (b) point-in-time sampling using the end-of-period observation (denoted PiT $x_t^{(3)}$) and (c) point-in-time sampling using the start-of-period observation (denoted PiT $x_{t-2/3}^{(3)}$). The sample size is $T = 40$ (top) or $T = 100$ (bottom). Frequencies for $s = 2$ represent power against a bivariate white noise. system.

For $T = 40$, power against a bivariate white noise system is lower when average sampling is employed (68-87%, depending on p) than when point-in-time sampling is used (81-95%, depending on p and whether the end- or start-of-period observation is concerned). When the sample size increases, this power is equal to 100% in almost all cases. The rejection frequencies corresponding to one common feature relationship, i.e., $s = 1$, depend on the temporal aggregation method chosen, p and T . For a given T , the frequencies increase with the lag length chosen for all three time aggregation approaches. Given p , the rejection frequencies also increase with the sample size (except for average sampling and $p = 6$). Finally, given p and T , the rejection frequencies are almost always largest when point-in-time sampling on the start-of-period

observation is performed, i.e., $\text{PiT}_t x_{t-2/3}^{(3)}$ (the exception being $p = 1$ and $T = 100$).

7 Conclusion

In this paper a new approach to common cycle analyses in multivariate time series processes is proposed when variables are sampled at mixed frequencies. This test allows users to examine commonalities in the business cycle of different countries when the frequency of observations is not the same. It also helps to detect co-movements when the economic indicators are not available at the same sampling frequency as in the applications given in this paper.

It is shown that in the presence of non-stationary variables, we need to account for two kinds of possible long-run relationships in a VAR; One stemming from the long-run relationship between the low- and high-frequency variables and several prespecified ones capturing the relationships between the high-frequency series themselves. A transformed VAR is presented, from which the set of variables is derived, that is used to construct an LR type test for the presence of common cyclical features. An empirical analysis and a small forecasting exercise with the quarterly GNP and monthly IPI illustrate this new approach. Finally, a Monte Carlo study is presented in order to investigate two issues: The frequencies with which multivariate information criteria detect different lag lengths and the rejection frequencies of the previously mentioned LR type common feature test. Using a mixed-frequency VAR as DGP, this analysis is done for our mixed-frequency models and several common (low-)frequency models obtained by temporal aggregation of the high-frequency variables.

From the empirical section as well as the Monte Carlo experiments it becomes clear that if the underlying data are of a mixed-frequency nature relying on mixed-frequency models instead of temporally aggregating the high-frequency variables may have advantageous effects. Firstly, we know that MIC consistently detect the correct lag length. Secondly, for the correct lag length, the common feature test behaves very well. And finally, common cycles, whose existence may be not visible when working with a low-frequency model, are detected and incorporated adequately in the model, which may lead to gains in terms of forecasting, for example. In other words, if there is useful information present in the high-frequency variables, there is some gain in making use of them and incorporate them adequately in a VAR model.

Some issues are not treated in this paper and are currently under study. A common trend/common cycle decomposition could be easily obtained given that the MF-SCCF can be seen as a special case of the PSCCF modeling. Naturally, the development of a strategy for higher dimensional systems is of importance. Finally, as mentioned in the Monte Carlo study, an extension of Cubadda et al. (2009) and Marcellino (1999) to a mixed-frequency transformed VAR is of interest in order to investigate and generalize the effects of temporal aggregation on the lag length selection and the analysis of common cycles in a mixed-frequency setup. A detailed study of the latter would also allow us to state what difference time aggregation of stock or flow variables makes in our framework. Furthermore, it could help answering the question whether mixed-frequency data help to detect the state of the business cycle earlier than with low-frequency data only.

References

- [1] Anderson, H. M., and Vahid, F. (2010), "VARs, Cointegration and Common Cycle Restrictions", Monash Econometrics and Business Statistics Working Papers 14/10, Monash University, Department of Econometrics and Business Statistics.
- [2] Andreou, E., Ghysels, E., and Kourtellis, R. (2010), "Regression Models with Mixed Sampling Frequencies", *Journal of Econometrics*, 158(2), 246-261.
- [3] Athanasopoulos, G., Guillén, O. T. de C., Issler, J. V., and Vahid, F. (2011), "Model Selection, Estimation and Forecasting in VAR Models with Short-Run and Long-Run Restrictions", *Journal of Econometrics*, 164, 116-129.
- [4] Clements, M. P., and A. B. Galvão (2007), "Marcoeconomic Forecasting with Mixed Frequency Data: Forecasting U.S. Output Growth", Working Papers 616, Queen Mary, University of London, School of Economics and Finance.
- [5] Clements, M. P., and A. B. Galvão (2009), "Forecasting U.S. Output Growth Using Leading Indicators: An Appraisal Using MIDAS Models", *Journal of Applied Econometrics*, 24(7):1187-1206.
- [6] Cox, D. R. (1981), "Statistical Analysis of Time Series: Some Recent Developments", *Scandinavian Journal of Statistics*, 8, 93-115.
- [7] Cubadda, G., and Hecq, A. (2001), "On Non-Contemporaneous Short-Run Co-Movements", *Economics Letters*, Elsevier, 73(3), 389-397.
- [8] Cubadda, G., Hecq, A. and F. Palm (2009), "Studying Co-movements in Large Multivariate Models Without Multivariate Modelling", *Journal of Econometrics*, 148(1), 25-35.
- [9] Engle, R. F., and Kozicki, S. (1993), "Testing for Common Features (with comments)", *Journal of Business and Economic Statistics*, 11, 369-395.
- [10] Forni, C., Ghysels, E., and Marcellino, M. (2012), "Mixed-Frequency Vector Autoregressive Models", This volume.
- [11] Ghysels, E. (2012), "Macroeconomics and the Reality of Mixed Frequency Data", mimeo University of North Carolina.
- [12] Gladyshev, E. (1961), "Periodically correlated random sequences", *Soviet Mathematics*, 2, 385-388.
- [13] Götz, T., Hecq A., and Urbain, J.-P. (2012a), "Forecasting Mixed-Frequency Time Series With ECM-MIDAS Models", Maastricht University METEOR Research Memorandum 12/012.

- [14] Götz, T., Hecq A., and Urbain, J.-P. (2012b), "Real-Time Forecast Density Combinations (Forecasting U.S. GDP Growth using Mixed-Frequency Data)", Maastricht University ME-TEOR Research Memorandum 12/021.
- [15] Hecq, A. (1998), "Does Seasonal Adjustment Induce Common Cycles?", *Economic Letters*, 59, 289-297.
- [16] Hecq, A., and Jacobs, J.P.A.M. (2009), "On the VAR-VECM Representation of Real Time Data", mimeo Maastricht University.
- [17] Hecq, A., Palm, F., and Urbain, J.-P. (2006), "Testing for Common Cyclical Features in VAR Models with Cointegration", *Journal of Econometrics*, 132 (1), 117-141.
- [18] Horvath, M. T. K., and Watson, M. W. (1995), "Testing for Cointegration when some of the Cointegrating Vectors are Prespecified", *Econometric Theory*, 11, 984-1014.
- [19] Lütkepohl, H. (2005), "New Introduction to Multiple Time Series Analysis", Berlin: Springer.
- [20] Marcellino, M. (1999), "Some Consequences of Temporal Aggregation in Empirical Analysis", *Journal of Business and Economic Statistics*, 17(1), 129-136.
- [21] Miller, J. I. (2011), "Cointegrating MIDAS Regressions and a MIDAS Test", Working Papers 1104, Department of Economics, University of Missouri.
- [22] Qian, H. (2010), "Vector Autoregression with Varied Frequency Data", MPRA Paper 34682.
- [23] Seong, B., Ahn, S. K., and Zadrozny, P. A. (2012), "Estimation of Vector Error Correction Models with Mixed-Frequency Data", *Journal of Time Series Analysis*, 1, 1-12.
- [24] Silvestrini, A., and Veredas, D. (2008), "Temporal Aggregation of Univariate and Multivariate Time Series Models: A Survey", *Journal of Economic Surveys*, 22(3), 458-497.
- [25] Vahid, F., and Engle, R. F. (1993), "Common Trends and Common Cycles", *Journal of Applied Econometrics*, 8(4), 341-360.
- [26] Vahid, F., and Issler, J. V. (2002), "The Importance of Common Cyclical Features in VAR Analysis: A Monte-Carlo Study", *Journal of Econometrics*, 109, 341-363.
- [27] Warne, A. (1993), "A Common Trends Model: Identification, Estimation and Inference", Seminar Paper No. 555, IIES, Stockholm University.
- [28] Zadrozny, P. (1988), "Gaussian Likelihood of Continuous-Time ARMAX Models When Data Are Stocks and Flows at Different Frequencies", *Econometric Theory*, 4(1), 108-124.

- [29] Zellner, A., and Palm, F. (1974), "Time Series Analysis and Simultaneous Equation Econometric Models", *Journal of Econometrics*, 2(1), 17-54.