Research Papers

ISSN 2610-931X

CEIS Tor Vergata

RESEARCH PAPER SERIES

Vol. 19, Issue 8, No. 525 – November 2021

Reduced Rank Regression Models in Economics and Finance

Gianluca Cubadda and Alain Hecq

Reduced Rank Regression Models in Economics and Finance

Gianluca Cubadda*

Università di Roma "Tor Vergata"

Alain Hecq[†]

Maastricht University

October 10, 2021[‡]

Abstract

This chapter surveys the importance of reduced rank regression techniques (RRR) for modelling economic and financial time series. We mainly focus on models that are capable to reproduce the presence of common dynamics among variables such as the serial correlation common feature and the multivariate autoregressive index models. Cointegration analysis, for which RRR plays a central role, is not discussed in this chapter as it deserves a specific treatment on its own. Instead, we show how to detect and model comovements in time series that are stationary or that have been stationarized after proper transformations. The motivations for the use of RRR in time series econometrics include dimension reductions which simplify complex dynamics and thus making interpretations easier, as well as pursuing efficiency gains in both estimation and prediction. Via the final equation representation, RRR also makes the nexus between multivariate time series and parsimonious marginal ARIMA models. The drawback of RRR, which is common to all the dimension reduction techniques, is that the underlying restrictions may be present or not in the data. We provide in this chapter a couple of empirical applications to illustrate concepts and methods.

Keywords: Reduced-rank regression, common features, vector autoregressive models, multivariate volatility models, dimension reduction.

*Università di Roma "Tor Vergata", Dipartimento di Economia e Finanza, Via Columbia 2, 00133 Roma, Italy. Email: gianluca.cubadda@uniroma2.it.

[†]Maastricht University, Department of Quantitative Economics, P.O.Box 616, 6200 MD Maastricht, The Netherlands. Email: a.hecq@maastrichtuniversity.nl.

[‡]This chapter has been prepared for the *Oxford Research Encyclopedia of Economics and Finance*. We thank two anonymous referees for helpful comments and corrections. The usual disclaimers apply.

1 Introduction

This chapter reviews, in an hopefully coherent manner, the main contributions in reduced rank regression models (RRR hereafter) for the multivariate analysis of economic and financial time series. Our aim is not to provide a complete historical survey of the RRR literature but instead to focus on, what we think, the most fruitful specifications of this modelling for applied econometricians are. We compare the different model representations and we illustrate them by applications on economic and financial time series. The goal of this chapter is to show to what extent RRR is helpful for applied researchers to detect comovements in times series and to reduce the dimensionality of the system as well as to improve the efficiency of estimators when RRR restrictions are taken in consideration.

We do not cover in details the statistical theory underlying RRR models, for which the reader is referred to, *inter alia*, Anderson (1984) and Reinsel and Velu (1998). We also focus in this chapter on specifications for time series that are stationary or that have been transformed to stationarity. Indeed, RRR has been extensively used for the analysis of cointegration among integrated variables (e.g. Johansen, 1996). We also do not deal with the asymptotic theory that is required when RRR is applied to non-stationary time series, which is a topic that deserves a complete treatment on its own.

The rest of this chapter is as follows. Section 2 motivates the use of RRR techniques and presents the most important specifications. Section 3 deals with the statistical inference. Section 4 goes more deeply into the properties of the most popular models, in particular common feature and autoregressive index models. In Section 5, several empirical applications illustrate our analyses.

1.1 Motivations for reduced rank regression models

Our starting point in this chapter is a multivariate time series model for a vector of n (weakly) stationary time series $Y_t = (y_{1t}, \dots, y_{nt})'$ such that

$$Y_t = \Pi X_t + \Gamma Z_t + \varepsilon_t, \quad t = 1 \dots T, \quad (1)$$

where X_t and Z_t are, respectively, vectors of k and g stationary time series, Π and Γ are, respectively, $n \times k$ and $n \times g$ coefficient matrices, and ε_t is a vector of n errors such that $E(\varepsilon_t | X_t, Z_t) = 0$. Both X_t and Z_t may contain (linear functions of) lags of Y_t , additional stochastic predictors and various deterministic functions of time such as an intercept, deterministic trends, trigonometric functions, and dummy variables.

We assume that $\text{rank}(\Gamma) = \min(n, g)$ and that it could be possibly equal to zero. Indeed, the main focus of this chapter will be on the matrix Π . Moreover, we assume for the moment that ε_t are i.i.d. innovations with $E(\varepsilon_t \varepsilon_t') = \Sigma$ (positive definite) and finite fourth moments. Gaussianity of the error terms will be invoked later for optimal statistical inference, in particular when we review the connection between the Canonical Correlation Analysis (CCA henceforth) and the Maximum Likelihood (ML) approach. Later, we will also relax the homoskedasticity assumption and consider the case when $E(\varepsilon_t \varepsilon_t') = \Sigma_t$. Finally, the key assumption of RRR is

$$\text{rank}(\Pi) = r < \min(n, k), \quad (2)$$

such that the matrix Π can be factorized as

$$\Pi = AB', \quad (3)$$

where A and B are, respectively, $n \times r$ and $k \times r$ full-rank matrices.¹ Notice that factorization (3) is clearly

¹A variant of this modelling consists in assuming that the matrix Γ has reduced rank as well. Statistical inference of such case can be found, *inter alia*, in Reinsel and Velu (1998) and Centoni *et al.* (2007).

not unique.

When condition (3) applies, Equation (1) defines, *stricto sensu*, a partial RRR model, whereas a full RRR model is obtained when Γ equal to zero as well. For the sake of simplicity, in the remainder of this chapter we will label *tout court* the following specification as the RRR model

$$Y_t = AB'X_t + \Gamma Z_t + \varepsilon_t, \quad t = 1, \dots, T. \quad (4)$$

Before we go into details about the several specifications of the RRR in econometrics, it is important to point out some differences between rank reduction techniques and other approaches such as principal component analyses (PCA) and other factor models. The main difference is that RRR imposes restrictions on the system generating the series, namely on the coefficient matrix Π , without losing any predictive accuracy. Indeed, one may insert between B' and X_t in (4) the following decomposition of the identity matrix

$$B(B'B)^{-1}B' + B_{\perp}(B'_{\perp}B_{\perp})^{-1}B'_{\perp} = I_n,$$

where $B_{\perp}$ is a full-rank $n \times (k - r)$ matrix such that $B'_{\perp}B = 0$, to see that series $B'_{\perp}X_t$ are irrelevant predictors of variables Y_t and hence can be safely ignored. In PCA, by construction, one encounters an information loss such that the first r principal components account for less than 100% of the total variation of the X_t variables for any $r < n$. The drawback of reduced rank restrictions is that they may not be present in the data, whereas it is always possible to extract principal components. In general, in the RRR approach the validity of the factorization (3) is tested and not imposed to the model if it is statistically rejected. As such, RRR models share the same aim as dimension reduction techniques or sparse methods, i.e. they impose proper restrictions on Π in order to simplify the system with virtually no costs in terms of prediction or explanation.

Being a dimension reduction technique, the first goal of RRR is consequently to reduce the number of parameters that have to be estimated. Indeed, under RRR there are at most

$$m = r(n + k - r) \quad (5)$$

parameters in Π that have to be estimated instead of $n \times k$ in the unrestricted specification. Let's take a numerical example with $n = 6$ variables, $r = 2$ and $k = 18$,² we have 108 parameters in the unrestricted matrix Π and only 44 parameters in the restricted product AB' . We subtract r^2 in m from the sum of the elements in A and B because the factorization (3) is defined up to a non-singular linear transformation. Indeed, given a $r \times r$ full-rank matrix F , we have

$$\Pi = AF^{-1}FB' \equiv A_*B'_*.$$

Determining the number of estimated parameters under reduced rank restrictions is important. It will be used to compute the penalty factor of information criteria to compare various specifications of the RRR. Similarly, the number of restrictions, namely $(n - r)(k - r)$, is the number of degrees of freedom to be used in likelihood ratio (LR henceforth) or Wald test statistics for the null hypothesis that restrictions (3) hold. We wrote above that there are "at most" m parameters because in some of the models that we will consider latter additional restrictions on A and B are imposed. For instance Cubadda and Hecq (2021) assume commonalities in the spaces generating A and B matrices, reducing further the number of estimated parameters and then m .

²This is the case for instance in a vector autoregressive model for 6 variables with 3 lags.

The first motivation for the use of RRR is of course linked to a dimension reduction. In principle, reducing the number of parameters should lead to more precise estimates of the parameters of interest when the rank restrictions are valid. At the empirical level, it is not guaranteed that imposing the RRR restrictions will lead to more precise inference and prediction. In particular, difficulties may arise when the unrestricted model has a large number of parameters w.r.t. the sample size and traditional methods, such as the ML approach, are used for statistical inference. Indeed, the favorable findings that Vahid and Issler (2002) obtained by simulations for a small vector auto-regressive model (VAR hereafter) using a canonical correlation analysis (CCA hereafter) were not replicated by other authors that considered medium or large VARs, see *inter alia* Cubadda and Hecq (2011) and Bernardini and Cubadda (2015). As we will see later in this chapter, RRR requires the use of non-conventional statistical methods when the system dimension is large.

The second reason of interest in RRR techniques is that they make the interpretation of multivariate time series models easier. Although imposing zero restrictions can be seen as a more natural way to reduce the number of parameters, parameters of reduced form models are often not endowed with a structural interpretation. Moreover, the factorization of Π allows for unravelling new insights on the presence of co-movements in economic and financial time series. As an example, it is well known that not only the presence cointegration is crucial in the analysis of spurious regressions but it is also informative for the interpretation of long run relationships between variables that are postulated by economic theory. Moreover, cointegration implies that a smaller number of common trends generates the non-stationary movements in the series. Similarly, RRR models can be used to investigate the presence of common features between stationary variables as defined by Engle and Kozicki (1993). Specifically, a feature is a given time series property (e.g. seasonality, autocorrelation, etc.) that is present in series Y_t and such feature is said to be common when there exist $n - r$ linear combinations of series Y_t that do not possess the feature anymore. For instance, a reduced-rank structure in a VAR model of the growth rates of different economies implies the presence of common cycles, which is turn informative about the degree of convergence of those economies (Vahid and Engle, 1993; Beine *et al.*, 2000; Hecq *et al.*, 2000). In finance, RRR models can be used to extract portfolios that have no time varying second moments although the volatilities of individual series are time varying (Hecq *et al.*, 2016). Moreover, the multivariate index autoregressive model (MAI hereafter), which is a VAR with a peculiar reduced-rank structure, can be used to extract common components in a similar manner as dynamic factor models (Reinsel, 1983).

The third reason why RRR models are practically relevant is for forecasting purposes. Multivariate time series models cannot be estimated with traditional statistical methods when the number of predictors is large w.r.t. the sample size. As a dimension reduction technique, RRR offers a possible way out to the problem of parameter proliferation that afflicts VAR models. Although empirical results that refer to small-scale models are somewhat mixed, RRR proved to be a valid alternative to Bayesian VARs and dynamic factor models for medium-large dimensional system, see *inter alia* Carriero *et al.* (2011) and Cubadda and Guardabascio (2019).

A fourth reason that has raised our own interest is that the presence common cyclical features restrictions, leading to reduced rank restrictions in the short-run dynamics, explains why we identify parsimonious univariate marginal ARIMA models in applied research, a paradox that the profession had difficulties to explain (see Cubadda *et al.*, 2008, 2009, Chevillon *et al.*, 2018; Hecq *et al.*, 2016).

Some of the most important specifications are briefly covered in the next subsection. A more detailed description of several models that we personally find the most interesting for empirical studies will be presented in Section 4.

1.2 A sketch of different models

1.2.1 Cointegration

This chapter considers RRR models for time series that are either stationary or stationarized after differencing. For instance we can use the first difference operator $\Delta = (1 - L)$, where L is the lag operator, or the seasonal differences $\Delta_s = (1 - L^s)$ for data that are observed s times in a year. In general, the choice of the operator to render series stationary is based on the outcome of some unit root tests.

It is a common practice in dynamic factor models to individually stationarize time series without paying attention to the presence of possible long-run relationships among variables. However, we cannot avoid to discuss RRR models without mentioning one of the most important example of common features, namely the presence of common stochastic trends and the related notion of cointegration (Engle and Granger, 1987). In this framework, we simplify the presentation by assuming that a VAR of finite order p generates n time series U_t that are integrated of order one (I(1) henceforth). Furthermore, we use the notation $Y_t = \Delta U_t$, $X_t = U_{t-1}$, $Z_t' = [D_t', \Delta U_{t-1}', \dots, \Delta U_{t-p+1}']$ where D_t' is a vector of deterministic elements. Under cointegration restrictions, the VAR model that generates series U_t can be rewritten in a vector error correction model (VECM henceforth), namely in a RRR model as (4) where A and B are full-rank $n \times r$ matrices such that $B'Y_{t-1}$, namely the error terms, are I(0) and A are their loadings. Since series U_t are non-stationary, a non-standard asymptotic theory is undertaken to evaluate the rank of Π , see *inter alia* Johansen (1996) and the references therein. Several extensions have been proposed in the literature such as multicointegration, seasonal cointegration, and fractional cointegration among others. A detailed description of the cointegration modelling is however beyond the scope of this survey.

1.2.2 Common deterministic features

Deterministic elements can also be subject to RRR commonalities. Assume that in Equation (1) we have that X_t is a vector of deterministic elements such as trigonometric functions, dummy variables and non-linear trends, whereas Z_t may include lags of Y_t as well as additional stochastic predictors. Examples of this modelling include the common deterministic seasonal features (Hylleberg and Engle, 1996; del Barrio Castro and Hecq, 2016; Hecq *et al.* 2012), co-breaking (Hendry, 1999), and co-trending (Bierens, 2000). We provide an example of common deterministic trends in the empirical section, this feature having been less investigated than others.

1.2.3 Common serial correlation

While the presence of cointegration between I(1) series reduces the rank of the long-run multiplier matrix, a common feature leads to a rank reduction in the matrix summarizing short-run dynamics. Most of this chapter emphasizes dynamic commonalities for a set of stationary time series. The starting point is the notion of the serial correlation common feature (SCCF henceforth) by Engle and Kozicki (1993). A stationary VAR(p) for series Y_t has the SCCF when in Equation (1) we have $X_t' = [Y_{t-1}', \dots, Y_{t-p}']$, Z_t is a vector of deterministic elements and

$$\Pi = A[B_1', \dots, B_p']. \quad (6)$$

where B_j is an $n \times r$ matrix for $j = 1, \dots, p$.

The set of restrictions in (6) implies that the coefficient matrix Π has a common left null space such that $A_\perp' \Pi = 0$ where $A_\perp' A = 0$. Several extensions have been proposed in the literature in order to relax the

restrictions (6) in various directions. In particular when SCCF has been analyzed together with cointegration, such as the notions of common cycles (Vahid and Engle, 1993), codependent cycles (Vahid and Engle, 1997), common cycles in seasonal non-stationary time series (Cubadda, 1999), the polynomial SCCF (Cubadda and Hecq, 2001), and weak form of SCCF (Hecq *et al.*, 2006).

1.2.4 Multivariate autoregressive index model

The specifications of the RRR that we have briefly illustrated so far are all special cases of common features as defined by Engle and Kozicki (1993). For instance, it is easy to see from (6) that the SCCF implies that the linear combinations $A'_\perp Y_t$ have no serial correlation although series Y_t are generated by a $\text{VAR}(p)$.

A RRR specification that does not belong to the common feature family (when $p > 1$ in the $\text{VAR}(p)$) is the multivariate autoregressive index model (MAI henceforth) as originally proposed by Reinsel (1988). Assume again that series Y_t are generated by stationary $\text{VAR}(p)$, hence $X_t = [Y'_{t-1}, \dots, Y'_{t-p}]'$ and Z_t is a vector of deterministic elements. The MAI requires that in Equation (1) we have

$$\Pi = [A_1, \dots, A_p](I_p \otimes \omega'), \quad (7)$$

where A_j is an $n \times r$ matrix for $j = 1, \dots, p$ and ω is a full-rank $n \times r$ matrix.

Notice that (7) can be equivalently rewritten as

$$\Pi = [A_1\omega', \dots, A_p\omega'],$$

which reveals that there is a clear link between the MAI and the popular dynamic factor model (Carriero *et al.*, 2016; Cubadda and Guardabascio, 2019) with the difference that no information is lost in the MAI and that specification might not exist and must be tested.

1.2.5 Common volatility

The RRR approach has also been used to investigate the second moments of financial time series. Among the various contributions, Hecq *et al.* (2016) and Engle and Marcucci (2016) consider a set the squared asset returns (or the residuals from a conditional mean model) and detect the presence of common volatilities. This means that although individual returns feature some time varying volatility, a combination of them, i.e. a portfolio, has got a non time varying conditional variance. Their study lies clearly within the multivariate GARCH framework. The investigation of Cubadda and Hecq (2017) differs from the previous ones in two aspects. First, they consider realized volatilities instead of squared returns within a vector heterogenous autoregressive process (VHAR). A VHAR is a $\text{VAR}(22)$ with restrictions on the parameters such that the only predictors are the first lags of the daily, weekly and monthly averages of series Y_t . Second, they extract the common components from a set of realized volatilities by means of a multivariate index model, focussing consequently on the right hand side properties of Π .

2 Statistical inference for reduced-rank regression models

In this section we review some of the statistical approaches that were proposed to identify and estimate the RRR model using time series data. We will distinguish the case where the matrices A and B are unconstrained, for which statistical inference on RRR can be obtained by CCA, from the case that restrictions apply either on A or B , such as in (7). In the latter case, a closed form solution of the ML estimator is not

always available. We will also mention the procedures that can be applied when the system dimension is large compared to the sample size.

2.1 Canonical correlation analysis

The estimation of reduced rank regression models involve the estimation of eigenvalues and eigenvectors of Π , usually by means of the canonical correlation procedure (Anderson, 1951). The first step of such procedure, that is called for when $\Gamma \neq 0$, requires to partial out the effect of Z_t both on Y_t and X_t . This goal is reached by computing the residuals $R_{0t} = Y'_t M_t$ and $R_{1t} = X'_t M_t$, where M_t is the t -th column of the projection matrix (residual maker)

$$M = I_T - Z(Z'Z)^{-1}Z',$$

$Y = [Y_1, \dots, Y_T]'$, $X = [X_1, \dots, X_T]'$, and $Z = [Z_1, \dots, Z_T]'$.

Next, we compute the cross-product matrices

$$S_{ij} = T^{-1} \sum_{i=1}^T R_{it} R'_{jt} \quad i = 0, 1; j = 0, 1.$$

Finally, we solve

$$|\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}| = 0$$

from which we extract the eigenvalues λ_i and the eigenvectors V_i for $i = 1, \dots, \min(n, k)$, where the former are sorted in non-increasing order. In practice, the eigenvalues may be conveniently obtained from the symmetric and semi-definite positive matrix $S_{11}^{-1/2} S_{10} S_{00}^{-1} S_{01} S_{11}^{-1/2}$ whose eigenvectors $\bar{V}_i$ are such that $V_i = S_{11}^{-1/2} \bar{V}_i$. In what follows, we denote such procedure as

$$\text{CanCorr}(Y_t, X_t | Z_t).$$

The key point of the asymptotics of $\text{CanCorr}(Y_t, X_t | Z_t)$ is that the smallest $n - r$ eigenvalues $\lambda_{r+1}, \dots, \lambda_n$ converge in probability to zero whereas the r eigenvectors $\hat{C} = [v_1, \dots, v_r]$ that are associated with $\lambda_1, \dots, \lambda_r$ are $\sqrt{T}$ -consistent estimator (up to a non-singular linear transformation) of the matrix B , see Anderson (1999).

Moreover, if the distribution of $[Y'_t, X'_t]'$ conditional on Z_t is Gaussian, CCA solves the ML problem for model (4). In particular, the LR test statistic for the null hypothesis that $\text{rank}(\Pi) = r$ versus the alternative that Π has full-rank is

$$LR(r) = -T \sum_{i=r+1}^{\min(n, k)} \ln(1 - \hat{\lambda}_i), \quad (8)$$

which is asymptotically distributed as

$$LR(r) \xrightarrow{d} \chi^2[(n - r)(k - r)].$$

whereas $\hat{C}$ is asymptotically normally distributed.

Finally, denoting $\hat{\Pi} = \hat{A} \hat{C}'$, where $\hat{A} = S_{10} \hat{C}$, $\hat{\Pi}$ is also asymptotically normally distributed with a covariance matrix that is smaller than the one in the unrestricted ordinary least square estimator (OLS) of Π . This last result does not depend on the normality assumption (Anderson, 2002). As previously mentioned, these asymptotics results do not hold if series Y_t posses unit roots. Readers are referred to Johansen (1996) for the case of cointegration.

Notice that, as an alternative to testing procedures, information criteria (IC) may be used to estimate r . Interestingly, IC for the problem at hand are linked to the statistic (8) by the following relation

$$IC(r) = -\frac{2}{T} \log \text{lik}_r + \frac{\rho(T, m)}{T},$$

where

$$\frac{2}{T} \log \text{lik}_r = \begin{cases} \ln |S_{00} - S_{01} S_{11}^{-1} S_{10}|, & \text{for } r = \min(n, k), \\ \frac{2}{T} \log \text{lik}_n - LR(r), & \text{for } r = 0, \dots, \min(n, k) - 1, \end{cases}$$

and the penalty function $\rho(T, m)$ is respectively $2m$, $m \ln \ln T$, and $m \ln T$ for the IC by Akaike, Hannan-Quinn, and Schwarz, see Vahid and Issler (2002) and the references therein. Hecq (2002) compares both approaches for the analysis of Latin American growth rates.

2.2 Switching algorithms

An alternative way of conducting inference on the RRR, which is particularly convenient when additional restrictions apply on either B or A , is using a switching algorithm as originally proposed by Boswijk (1995) for cointegration analysis. Basically, a switching algorithm is an iterative procedure that aims at increasing the likelihood in each step by alternating between the maximization of two different object functions: the likelihood of A , Γ and Σ conditional on B , and the likelihood of B conditional on A , Γ and Σ .

In particular, under the assumption of Gaussianity the switching algorithm for estimating model (4) goes as follows:

1. In view of Equation (4) and given (initial) estimates of B , estimate A , Γ and Σ by OLS.
2. Premultiply by $\Sigma^{-1/2}$ and apply the Vec operator to both the sides of Equation (4), then use the property $\text{Vec}(ABC) = (C' \otimes A) \text{Vec}(B)$ to get

$$\Sigma^{-1/2}(Y_t - \Gamma Z_t) = (X_t' \otimes \Sigma^{-1/2} A) \text{Vec}(B') + \Sigma^{-1/2} \varepsilon_t$$

Given the previously obtained estimates of A , Γ and Σ , estimate $\text{Vec}(B')$ with OLS on the above model.

3. Repeat Steps 1 and 2 till numerical convergence occurs.³

The main merit of switching algorithms is that they allow for easily imposing over-identifying restrictions or regularization schemes in the estimation of B or A . For instance, Cubadda *et al.* (2017) proposed a switching algorithm for endowing a multivariate model for realized volatilities with the index structure in (7), whereas Cubadda and Guardabascio (2019) offered switching algorithms for the regularized estimation of the SCCF and MAI models that are augmented with individual AR components for each element of Y_t .

A classical alternative to switching algorithms is the Newton-Raphson method used in Reinsel and Ahn (1992), which explicitly take into account the correlation structure of the estimates of A and B .

2.3 Reduced rank regression under heteroskedasticity

We have assumed that the errors ε_t are homoscedastic; we now look at the impact of time varying volatility when assessing the presence of comovements. We review to what extent the presence of heteroskedasticity

³A general proof of the convergence of this family of iterative procedures is given by Oberhofer and Kmenta (1974).

modifies the behavior of the tests statistics covered so far in this chapter. For instance, to what extent can we accurately detect the presence of common dynamics in series like interest rates, inflation rates, exchange rates or stock returns, namely most macro finance data, when variables are characterized by time varying second moments.

To the best of our knowledge, Beine and Hecq (1999) is one of the first paper that investigates, using Monte Carlo simulations, the impact of GARCH type errors when testing for reduced rank in stationary multivariate time series.⁴ They show that testing for reduced rank restrictions in multivariate moving average models (a framework called codependence by Gouriéroux and Peaucelle, 1993) is heavily affected by the presence of heteroskedastic errors. Hecq (2000) reaches similar conclusions for the scalar component model of Tiao and Tsay (1989), namely a VAR-moving average model with potential reduced rank restrictions on both the autoregressive and the moving average parts. Both Beine and Hecq (1999) and Hecq (2000) point out that time varying second moments also lead to an overestimation of the detection of lag orders of dynamic models using conventional strategies, hence also reducing the power of reduced rank tests. These findings extend to the multivariate case those obtained by Diebold (1986) and Hecq (1996) for univariate time series. GARCH errors are then responsible for large size distortions in the sense that traditional critical values are too low for correctly rejecting the null. As expected, the distortions observed increase when volatility clusters are large and with the sample size as the occurrence of large volatile events increases. However, neither Beine and Hecq (1999) nor Hecq (2000) propose an alternative modeling. We next review some possible robust solutions that have been developed in the literature.

Let us first focus on the estimation of the matrix $A'_\perp$ when $X'_t = [Y'_{t-1}, \dots, Y'_{t-p}]$, Z_t is a vector of deterministic elements and restrictions (6) apply. From a generalize method of moments (GMM) perspective, the SCCF null hypothesis consists of an orthogonality condition between a combination of variables Y_t and their p lags X_t . Consequently the use of the GMM and the associated orthogonality test is quite intuitive. Candelon *et al.* (2005) introduce robust parametric and non parametric procedures to test for serial correlation common features in the presence of time varying second moments. We only describe their robust GMM approach that is more closely related to the material of this chapter. In a nutshell, the idea behind their robust GMM approach is simply to extend the single equation instrumental variable estimator (IV) using heteroskedastic consistent standard errors (HCSE) proposed by White (1982).

Assume for the moment that $n - r = 1$, hence $A'_\perp$ is a n -vector, and write $A'_\perp = [1, -\theta']'$ and $Y_t = [y_{1t}, W_t]'$. In the presence of heteroskedastic errors, Candelon *et al.* (2005) consider the following GMM estimator of θ

$$\hat{\theta} = (W'MX(X'MSMX)^{-1}X'MW)^{-1}W'MX(X'MSMX)^{-1}X'My_1$$

where $y_1 = [y_{1,1}, \dots, y_{1,T}]'$, $W = [W_1, \dots, W_T]'$, $S = \text{diag}(e_1^2, \dots, e_T^2)$, and $e_t = y_{1t} - \tilde{\theta}'W_t$, where $\tilde{\theta}$ is the homoscedastic version of the estimator above, i.e. with $S = I_T$. A test for the existence of a single SCCF vector can then be via the usual overidentification test statistic

$$J(\hat{\theta}) = e^{*'}MX(X'MSMX)^{-1}X'Me^*$$

where $e^* = [e_1^*, \dots, e_T^*]'$, and $e_t^* = y_{1t} - \hat{\theta}'W_t$. Under the null hypothesis we have $J(\hat{\theta}) \xrightarrow{d} \chi^2[n(p-1)+1]$.

Candelon *et al.* (2005) show in their simulation study that the sizes of both the non-robust GMM and the canonical correlation procedures increase dramatically with the impact of yesterday's news parameter in the GARCH as well as with the number of observations.⁵ Notice that the single robust GMM approach can

⁴There exists a large literature on the impact of GARCH innovations on cointegration analysis that we do not cover.

⁵Hecq and Issler (2012) have also implemented a Newey and West (1987) correction in constructing S .

be easily extended to the multivariate robust case using the normalization

$$A'_{\perp} = [I_{n-r}, -\Theta]$$

where Θ is full rank $(n-r) \times (n-r)$ matrix. However, a limit of the multivariate GMM approach is that it is still a regression strategy that is dependent on the choice of left hand side variables, which is generally arbitrary. Hence, we would favor a multivariate robust counterpart to the canonical correlation approach.

With that aim in mind, Hecq and Issler (2012) propose to modify the likelihood ratio test in two aspects. First they use a Wald approach introduced in Christensen, Hurn and Pagan (2011). Having reformulated the SCCF restrictions as

$$R\text{Vec}(\Pi') = 0_{(snp) \times 1}$$

where $R = A'_{\perp} \otimes I_{np}$, the latter authors propose to substitute $A'_{\perp}$ with its estimate $\hat{A}_{\perp}$ that is obtained by CCA. Under homoscedasticity, they show that such Wald test is asymptotically equivalent to the LR one. Then Hecq and Issler (2012) robustify the considered Wald test using the multivariate extension of the White's HCSE proposed in Ravikumar *et al.* (2000) for system of seemingly unrelated regressions. Indeed, HCSE does not modify the estimated parameters nor the likelihood and hence a Wald test is appropriate. The Monte Carlo simulations in Hecq and Issler (2012) show that the proposed strategy is as good as a single equation robust GMM for $n-r=1$, and that it performs well to detect multiple SCCF relationships.

Alternatively, Kleibergen and Paap (2006) provided a general approach for testing restrictions (6) that allows for heteroscedastic errors in model (4). Their method requires as a first step to estimate unrestrictedly the matrix Π by an estimator $\hat{\Pi}$ that has an asymptotic normal distribution, and then to compute the singular value decomposition $\hat{\Pi} = UDV$, where D is the $n \times k$ rectangular diagonal matrix of the singular values $d_1, \dots, d_n$ which are sorted in non-increasing order, and finally obtaining a test statistic for the null hypothesis (6) that is based on proper transformation of the $(n-r)$ smallest singular values $d_{r+1}, \dots, d_n$. The key point is that the approach is general enough to allow for the use of an estimator of Π such that the asymptotic variance matrix of $\text{Vec}(\hat{\Pi})$ has not a Kronecker structure as in the case of heteroscedasticity. They show by simulation that their test, in combination with an estimator $\hat{\Pi}$ that explicitly takes into account the presence of heteroscedastic errors, outperforms the one based on the standard CCA approach.

2.4 Inference for high-dimensional models

The previously illustrated procedures are based on the standard asymptotic framework where the sample size T diverges and the system dimension n is fixed. A relevant part of modern time series econometrics focuses instead on the high-dimensional environment where both T and n are large, with the former possibly smaller than the latter. It is clear that in such a framework standard estimation procedures break down and new methods are called for.

Focusing on the SCCF case, Lam *et al.* (2011) proposed a simple non-parametric estimator of the matrix A that is consistent when T and even n diverge. The basic assumptions are that series Y_t are stationary, $\Gamma = 0$,⁶ $E(Y_t|F_{t-1}) = AB'X_t$, where F_t is the natural filtration of the process Y_t at time t , and, without loss of generality that $A'A = I_r$. Then we can use the identity $AA' + A_{\perp}A'_{\perp} = I_n$ in order to decompose series Y_t as

$$Y_t = Af_t + A_{\perp}\eta_t, \tag{9}$$

where $f_t = A'Y_t$, $\eta_t = A'_{\perp}\varepsilon_t$ and $A_{\perp}$ is a full-rank $n \times (n-r)$ matrix such that $A'_{\perp}A = 0$ and $A'_{\perp}A_{\perp} = I_{n-r}$.

⁶See Chang *et al.* (2015) for an extension of the method to the case that $\Gamma \neq 0$.

We see from (9) that $A'_{\perp} Y_t = \eta_t$ is a $(n - r)$ dimensional innovation process. Hence, the matrix A lies in the space generated by the eigenvectors associated with the r non-zero eigenvalues of the symmetric and semi-positive definite matrix

$$M = \sum_{j=1}^{p_0} \Sigma_y(j) \Sigma_y(j)',$$

where p_0 is a positive integer, where $\Sigma_y(j)$ is the autocovariance matrix of series Y_t at lag j . Let us indicate with $\hat{V}_q$ the matrix formed by the eigenvectors associated with the q ($\leq n$) largest eigenvalues of the sample estimate of M . Under some regularity conditions, $\hat{V}_r$ is a $\sqrt{T}$ -consistent estimator of A (up to an orthonormal transformation) when r is fixed, $n, T \rightarrow \infty$, and $a'_i a_i = O(n)$ for $i = 1, \dots, r$, where $A = [a_1, \dots, a_r]$. Moreover, Lam and Yao (2012) proved that a consistent estimator of r is provided by

$$\hat{r} = \arg \min_{i=1, \dots, R} \left\{ \hat{\lambda}_{i+1} / \hat{\lambda}_i \right\}, \quad (10)$$

where R is a constant such that $r < R < n$ and $\hat{\lambda}_i$ is the i -th largest eigenvalue of matrix $\hat{M}$.

Remarkably, Equation (9) shows that the factors f_t are responsible for the whole dynamic of series Y_t . Under the assumption that variables Y_t are generated by a large stationary VAR, Cubadda and Hecq (2021) provide the conditions under which the factors f_t are generated by a (small-scale) VAR as well as the related statistical methods.

In a Bayesian setting, Carriero *et al.* (2011) proposed a straightforward method for imposing the SCCF restrictions (6) in large VARs. Assuming $X'_t = [Y'_{t-1}, \dots, Y'_{t-p}]$ and that Z_t is a vector of deterministic elements, the suggested approach consists in first estimating the posterior mean $\bar{\Pi}$ of the matrix Π by a proper version of the Minnesota prior (Kadiyala and Karlsson, 1997), then computing the singular value decomposition $\bar{\Pi} = UDV$, where D is the $n \times k$ rectangular diagonal matrix of the singular values $d_1, \dots, d_n$ which are sorted in non-increasing order, and finally obtaining a reduced rank approximation of the posterior mean as

$$\bar{\Pi}_r = [u_1, \dots, u_r] D_r [v_1, \dots, v_r]$$

where $D_r = \text{diag}[d_1, \dots, d_r]$, and u_i, v_i are the vectors associated with d_i for $i = 1, \dots, n$. Carriero *et al.* (2011) empirically show that the proposed method compares well with popular approaches in macroeconomic forecasting.

Other approaches to tackle the dimensionality issue in RRR models include Cubadda and Hecq (2011), who proposed a multivariate partial least squares approach to test for the presence of SCCF, Bernardini and Cubadda (2015), who offered a regularized version of the CCA approach for stationary VARs, and Carriero *et al.* (2016), who developed Bayesian inference for the MAI modelling.

3 Reduced-rank structures for economic and financial modelling

In this section we review in more details the most common specifications of the RRR approach for the analysis of economic and financial variables. The first two subsections follow and update Centoni and Cubadda (2015) in reviewing models that are mainly applied in macroeconometrics whereas the last subsection focuses on the use of RRR in financial econometrics.

3.1 Common cyclical features

Let us assume that variables U_t are generated by a VAR(p) model and that they are cointegrated of order (1, 1). Hence, they admit the following vector error correction model (VECM) representation:

$$\Gamma(L)\Delta U_t = AB'U_{t-1} + \varepsilon_t \quad (11)$$

where $\Gamma(L) = I_n - \sum_{i=1}^{p-1} \Gamma_i L^i$ such that the solutions of $\det[\Gamma(z)] = 0$ lie outside the unit circle and $B'U_t$ is I(0). Since ΔU_t are I(0), they admit the following Wold representation:

$$\Delta U_t = \Psi(L)\varepsilon_t$$

where $\Psi(L) = I_n + \sum_{i=1}^{\infty} \Psi_i L^i$ is such that $\sum_{j=1}^{\infty} j |\Psi_j| < \infty$. By expanding $\Psi(L)$ on 1 and integrating both sides of the above equation, we get the multivariate Beveridge-Nelson representation (BN):

$$U_t = U_0 + \tau_t + \kappa_t$$

where $\tau_t = \Psi(1) \sum_{j=0}^{t-1} \varepsilon_{t-j}$, $\kappa_t = \sum_{j=0}^{t-1} \tilde{\Psi}_j \varepsilon_{t-j}$, and $\tilde{\Psi}_i = -\sum_{j>i} \Psi_j$ for all i .

We know from the Engle-Granger representation theorem (Engle and Granger, 1987; Johansen, 1996) that $B'\Psi(1) = 0$ and thus $B'\tau_t = 0$, series U_t share $(n - r)$ common stochastic trends. However, detrended macroeconomic time series typically display clear evidence of comovements (Lucas, 1977), thus suggesting that common trends are not the only source of commonalities in the economy. This leaves open the possibility that there exist linear combinations of cyclical series that are not cyclical. In order to expand this argument, let us concentrate on κ_t , namely the stochastic cycles of series U_t . While the presence of cointegration implies reduced-rank restrictions on the VECM parameters that are responsible for the presence of common trends, the common cyclical feature (CCF henceforth) approach is instead concerned with reduced-rank restrictions on the VECM parameters that induce commonalities among the cycles κ_t . Remarkably, timing is important when defining a notion of short-run comovements. Hence, we have alternative notions of CCFs that differ according to the degree of synchronicity that they impose on the cycles κ_t .

The seminal notion of CCF is the one of SCCF (Engle and Kozicki, 1993). In particular, series ΔU_t have s ($s < n$) SSCFs iff there exists an $n \times s$ matrix δ_S with full column rank such that the VECM (11) can be rewritten into the following RRR model

$$\Delta U_t = \delta_{S\perp} \omega'_S (U'_{t-1} B, \Delta U'_{t-1}, \dots, \Delta U'_{t-p+1})' + \varepsilon_t$$

where ω_S is an $(np - n + r) \times (n - s)$ matrix with full column rank. Since $\delta'_S \Delta U_t = \delta'_S \varepsilon_t$, there exist s linear combinations of series ΔU_t that are innovation processes. Since $\delta'_S \kappa_t = 0$, there exist $(n - s)$ common cycles in series U_t that are perfectly synchronized (Vahid and Engle, 1993).

Cubadda and Hecq (2001) proposed the notion of polynomial serial correlation common features (PSCCF) as a measure of non-contemporaneous cyclical comovements. Non-synchronous common cycles arises, for example, in economic model of consumption with several types of consumer goods as, *e.g.*, in Vahid and Engle (1997) and Schleicher (2007); it is shown that the maximization problem of the representative agent implies an unsynchronized common cycle in the consumer goods vector. We say that series ΔU_t have s PSCCF's iff there exists an $n \times s$ matrix δ_P with full column rank such that $\delta'_P \Gamma_1 \neq 0$, and thus the VECM (11) can be rewritten into the following partial RRR model

$$\Delta U_t = \Gamma_1 \Delta U_{t-1} + \delta_{P\perp} \omega'_P (\Delta U'_{t-2}, \dots, \Delta U'_{t-p+1}, U'_{t-1} B)' + \varepsilon_t$$

where ω_P is an $(np - 2n + r) \times (n - s)$ matrix with full column rank. Remarkably, the existence of PSCCF requires that there exists a first-order polynomial matrix

$$\delta(L) = \delta_P - \Gamma'_1 \delta_P L$$

such that $\delta(L)' \Delta U_t = \delta'_P \varepsilon_t$, and $\delta(L)' \kappa_t = \delta'_1 \Psi(1) \varepsilon_t$. Hence, the PSCCF relationships cancel the dependence from the past of both the first differences and the cycles of series U_t .

In the previous definitions of common cyclical features, the number of CCF relationships, s , cannot exceed the number of common trends, $(n - r)$. In order to remove this restrictions, Hecq *et al.* (2006) proposed the notion of weak form of SCCF (WF henceforth): series ΔU_t have s WF's iff there exists an $n \times s$ matrix δ_W with full column rank such that $\delta'_W A \neq 0$, and the VECM in (11) can be rewritten into the following partial RRR model

$$\Delta U_t = AB'U_{t-1} + \delta_W \omega'_W (\Delta U'_{t-1}, \dots, \Delta U'_{t-p+1})' + \varepsilon_t$$

where ω_W is an $(np - n) \times (n - s)$ matrix with full column rank. Cubadda (2007) shows that WF requires the existence of a first-order polynomial matrix

$$\delta_W(L) \equiv \delta_W - (BA' + I_n) \delta_W L$$

such that $\delta_W(L)' U_t = \delta'_W \varepsilon_t$, and $\delta_W(L)' \kappa_t = \delta'_W (I_n - \Psi(1)) \varepsilon_t$. Hence, the same WF relationships cancel the dependence from the past of both the levels and the cycles of series U_t .

A limitations of the previous models of CCFs is that they cannot handle the possible coexistence of differing types of reduced-rank restrictions within the same model. In order to overcome this limitation, Cubadda (2007) introduced the notion of weak form of PSCCF (WFP henceforth), which encompasses most of the existing formulations: series ΔU_t have s WFP's iff there exists an $n \times s$ matrix δ_F with full column rank such that $\delta'_F A \neq 0$, $\delta'_F \Gamma_1 \neq 0$, and the VECM in (11) can be rewritten into the following partial RRR model

$$\Delta U_t = AB'U_{t-1} + \Gamma_1 \Delta U_{t-1} + \delta_F \omega'_F (\Delta U'_{t-2}, \dots, \Delta U'_{t-p+1})' + \varepsilon_t$$

where ω_F is an $(np - 2n) \times (n - s)$ matrix with full column rank. The WFP requires the existence of a second-order polynomial matrix

$$\delta_F(L) \equiv \delta_F - (BA' + I_n + \Gamma'_1) \delta_F L + \Gamma'_1 \delta_F L^2$$

such that $\delta_F(L)' U_t = \delta'_F \varepsilon_t$ and $\delta_F(L)' \kappa_t \sim \text{VMA}(1)$. Hence, the polynomial matrix $\delta_F(L)$ transforms the BN cycles κ_t into a process with shorter memory.

Cubadda *et al.* (2008, 2009) focused on the implications of the existence of CCFs for the univariate autoregressive integrated moving average (ARIMA) representations of elements of U_t . As it is well known, each series generated by a VAR process follows an individual ARIMA process. However, the VAR models that are typically used in macroeconomic analysis would imply highly non parsimonious ARIMA models for individual time series, whereas low order ARIMA models are empirically appropriate in most cases. This is the so-called "autoregressivity paradox". Remarkably, Cubadda *et al.* (2009) show that when the VAR or VECM model is endowed with CCF restrictions, the individual ARMA orders are significantly reduced, becoming much closer to those that are commonly employed in empirical applications. Hence, the existence of CCF's provides a possible, economically meaningful, solution of the autoregressivity paradox.

The CCF analysis was extended even to series having different forms of non-stationarity than I(1)-ness. In particular, Cubadda (1999, 2001) explored the presence of common cycles in variables that are also integrated

at (a subset of) the seasonal frequencies, whereas Paruolo (2006) focused on the case of I(2) systems. Franchi and Paruolo (2011) offered a comprehensive theoretical analysis of the conditions of existence of the various form of CCF's and of the characterization of the CCF relations in I(0), I(1) and I(2) systems.

Bulding on Ahn and Reinsel (1989), Ahn (1997) proposed a variant of the SCCF such that coefficient matrices of the lags of ΔU_t have reduced ranks that are nested each other: $\text{Range}(\Gamma_j) \supset \text{Range}(\Gamma_{j+1})$ for $j = 1, \dots, p-2$. Such nested structure provides a very general statistical model since it encompasses both the SCCF and its polynomial extensions as particular cases. However, its interpretation in terms of common short-run components is more involved.⁷

Goetz *et al.* (2016) show how RRR restrictions can help in reducing the number of parameters when testing for Granger-causality in large mixed-frequency VARs with both monthly low frequency and daily high frequency series. Goetz *et al.* (2013) detect the presence of common cycles in mixed frequency VARs when the number of high frequency series is not too large compared to the low frequency variable.

Recently, Cubadda and Hecq (2019) investigated the presence of CCFs in pure non-causal VARs, i.e. in multivariate time series models where variables Y_t are linear functions of their leads plus an error term. A complication in the statistical analysis of those models is that each non-causal VAR admits a so-called pseudo-causal representation, i.e. a traditional VAR model where the error term is a white noise but not a innovation process. Since Gaussian ML cannot discriminate between the genuine non-causal representation and the pseudo representation of the series, it is required to assume that errors are not normal in order to exploit moments higher than two for statistical inference. Interestingly, Cubadda and Hecq (2019) show that if the non-causal VAR has the SCCF, the pseudo-causal VAR has not and *vice versa*. Hence, CCA can be applied, where either lags or leads are included in X_t , to determine whether SCCF is present in either the backward- or forward-looking dynamics of the series.

3.2 VAR models with an index structure

Another modelling that is also endowed with a reduced-rank structure is the Multivariate Autoregressive Index model (MAI henceforth) as originally proposed by Reinsel (1983). Assuming that series Y_t are generated by a stationary VAR(p) model and that restrictions (7) apply, the basic version of the MAI reads

$$Y_t = \sum_{j=1}^p A_j \omega' Y_{t-j} + \varepsilon_t \quad (12)$$

The linear combinations $f_t = \omega' Y_t$ are called the indexes and, interestingly, they follow a VAR(p) process as well. Indeed, if we premultiply both sides of Equation (12) by the matrix ω' we get

$$f_t = \sum_{j=1}^p \omega' A_j f_{t-j} + \omega' \varepsilon_t, \quad (13)$$

whereas linear combinations of series generated by an unrestricted VAR(p) model generally follow a VARMA process, see Cubadda *et al.* (2009) and the references therein.

Lately, the MAI has gained some popularity because it allows to rewrite a (possibly large) VAR model in an analogous way to the popular dynamic factor model (DFM, surveyed in Lippi, 2019). In order to see this, invert the polynomial matrix in (12) to get the Wold representation of series Y_t

$$Y_t = \Psi(L)\varepsilon_t,$$

⁷Ahn and Reinsel (1989) interpreted the nested reduced-rank structure in terms of a reduced number of channels (w.r.t. n) through which past information is passed onto the current observation.

then use the decomposition of the identity matrix as in Centoni and Cubadda (2003)

$$\Sigma\omega(\omega'\Sigma\omega)^{-1}\omega' + \omega_{\perp}(\omega'_{\perp}\Sigma^{-1}\omega_{\perp})^{-1}\omega'_{\perp}\Sigma^{-1} = I_n$$

to obtain the decomposition

$$Y_t = \chi_t + \iota_t \quad (14)$$

where

$$\begin{aligned} \chi_t &= \Psi(L)\Sigma\omega(\omega'\Sigma\omega)^{-1}\varepsilon_t^{\chi}, \\ \iota_t &= \Psi(L)\omega_{\perp}(\omega'_{\perp}\Sigma^{-1}\omega_{\perp})^{-1}\varepsilon_t^{\iota}, \end{aligned}$$

$\varepsilon_t^{\chi} = \omega'\varepsilon_t$, $\varepsilon_t^{\iota} = \omega'_{\perp}\Sigma^{-1}\varepsilon_t$. Since $E(\varepsilon_t^{\chi'}\varepsilon_t^{\iota'}) = 0$, the components χ_t and ι_t are uncorrelated at all lags and leads. Moreover, we see from (12) that shocks ε_t^{χ} are those hitting the indexes f_t . Hence, we can refer to ε_t^{χ} as the common shocks of series Y_t and, consequently, to χ_t as the common components of series Y_t .

Carriero *et al.* (2016) proposed a Bayesian method to estimate the MAI and then used decomposition (14) to perform a structural analysis where the structural errors are linear transformation of the common shocks ε_t^{χ} only.

Cubadda and Guardabascio (2019) extended the basic MAI model by allowing for individual AR structures for each element of Y_t . In particular the proposed model, which they label as the multivariate index-augmented autoregressive model (MIAAR), reads

$$Y_t = \sum_{j=1}^p D_j Y_{t-j} + \sum_{j=1}^p A_j \omega' Y_{t-j} + \varepsilon_t, \quad (15)$$

D_j is a $N \times N$ diagonal matrix for $h = 1, \dots, p$. Remarkably, the mathematical formulation of the MIAAR model is such that series are forecasted in a similar manner as by the DFM. Cubadda and Guardabascio (2019) offer a switching algorithm, along with its regularized version, to estimate the parameters of model (15), and they empirically show that the MIAAR compares well with popular approaches in macroeconomic forecasting.

3.3 Reduced rank models for second moments

This subsection does not see time varying volatility as a misspecification one has to robustify for but instead a feature of interest *per se*. Instead, we present model for the detection of commonalities in second moments. Many models have been proposed to proxy volatilities, that are in essence latent variables. If one takes the example of stock price series, one can test for reduced rank restrictions on squared returns using linearized multivariate version of the generalized auto-regressive conditional heteroskedasticity (GARCH) model. Alternatively, one can directly estimate multivariate VAR models on realized volatility series built from intraday data. We cover models such as common ARCH features modelings when there exists a common left null space in the parameters of the multivariate model generating the squared returns. We also look at multivariate index models for realized (co-)volatilities.

3.3.1 Common GARCH features

Let us denote $\varepsilon_t = H_t^{1/2}z_t$, $t = 1, \dots, T$, the $n \times 1$ vector of excess returns of financial assets observed at the time period $t = 1 \dots T$ with $z_t \sim iid.(0, I_n)$. Consequently we have $\varepsilon_t | \Omega_{t-1} \sim D(0, H_t)$ with Ω_{t-1}

the past information set and H_t the conditional covariance matrix of the n assets that is measurable with respect to Ω_{t-1} ; $D(.,.)$ is some arbitrary multivariate distribution. Using half-vectorization operators for the $\text{Vech}(H_t)$, let us denote $v_t = \text{Vech}(\varepsilon_t \varepsilon_t') - \text{Vech}(H_t)$ which by definition is a martingale difference sequence with respect to the past of $\text{Vech}(\varepsilon_t \varepsilon_t')$. A multivariate GARCH(0, 1) can be written as a VAR(1) for observed squared returns and covariances. For instance a bivariate model with $\text{Vech}(H_t) = \{h_{11t}, h_{12t}, h_{22t}\}'$ and $\text{Vech}(\varepsilon_t \varepsilon_t') = \{\varepsilon_{1t}^2, \varepsilon_{1t}\varepsilon_{2t}, \varepsilon_{2t}^2\}'$ gives

$$\begin{pmatrix} \varepsilon_{1t}^2 \\ \varepsilon_{1t}\varepsilon_{2t} \\ \varepsilon_{2t}^2 \end{pmatrix} = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} + \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix} \begin{pmatrix} \varepsilon_{1t-1}^2 \\ \varepsilon_{1t-1}\varepsilon_{2t-1} \\ \varepsilon_{2t-1}^2 \end{pmatrix} + \begin{pmatrix} v_{1t} \\ v_{2t} \\ v_{3t} \end{pmatrix}, \quad (16)$$

where usual non-negativity and stationarity restrictions on the parameters are assumed to be met. For n assets and q lags, we have a $N = n(n+1)/2$ dimensional vector, such that

$$A(L)\epsilon_t = \omega + v_t, \quad (17)$$

where $A(L) = (I - A_1L - \dots - A_qL^q)$, A_i is a $n \times n$ coefficient matrix for $i = 1, \dots, q$, ω and v_t are n -dimensional vectors. Testing for common GARCH feature (in the sense of Engle and Kozicki, 1993) amounts to look towards s directions $\delta'\varepsilon_t$ such that the $\delta'\varepsilon_t$ combinations of asset returns are conditionally homoskedastic, namely such that $\delta'\varepsilon_t|\Omega_{t-1} \sim D(0, C)$ where C does not depend on t . There exist s linear combinations of ε_t that have time-invariant conditional distributions. The remaining $n - s$ combinations generate the time-varying volatility of the system.

Engle and Marcucci (2006) use canonical correlation between pairs of (log) squared returns series $\text{Vecd}(\varepsilon_t \varepsilon_t') = \{\varepsilon_{1t}^2, \varepsilon_{2t}^2, \dots, \varepsilon_{nt}^2\}'$ to investigate comovements. They estimate covariances separately and reconstruct the variance-covariance matrix using the dynamic conditional correlation specification. This is the pure variance model. Hecq *et al.* (2016) compare the pure variance model with specifications enriched with time varying covariances and propose restrictions in three VAR specifications, the last one with variances only (Vecd operator) being similar to Engle and Marcucci (2006).

$$\begin{aligned} \mathcal{M}_1 : \quad & \text{Vech}(\varepsilon_t \varepsilon_t') = \gamma_0 + A_1 \text{Vech}(\varepsilon_{t-1} \varepsilon_{t-1}') + \dots + A_q \text{Vech}(\varepsilon_{t-p} \varepsilon_{t-p}') + v_t, \\ \mathcal{M}_2 : \quad & \text{Vecd}(\varepsilon_t \varepsilon_t') = \tilde{\gamma}_0 + \tilde{A}_1 \text{Vecd}(\varepsilon_{t-1} \varepsilon_{t-1}') + \dots + \tilde{A}_q \text{Vecd}(\varepsilon_{t-p} \varepsilon_{t-p}') + \tilde{v}_t, \\ \mathcal{M}_3 : \quad & \text{Vecd}(\varepsilon_t \varepsilon_t') = \check{\gamma}_0 + \check{A}_1 \text{Vecd}(\varepsilon_{t-1} \varepsilon_{t-1}') + \dots + \check{A}_q \text{Vecd}(\varepsilon_{t-p} \varepsilon_{t-p}') + \check{v}_t, \end{aligned}$$

where v_t and $\tilde{v}_t$ are martingale differences of the Wold or VAR representation of $\text{Vech}(\varepsilon_t \varepsilon_t')$ and $\text{Vecd}(\varepsilon_t \varepsilon_t')$, and $\check{v}_t$ is a martingale difference if the DGP is a pure variance model. As they are martingale differences, they might exhibit heteroskedasticity but not autocorrelation. Contrarily to a model with variances only, the drawback of the first two VARs is that log transformations are not possible, which keeps the errors non-Gaussian.

Simulation studies in Hecq *et al.* (2016) suggest that standard canonical correlation tests (in the spirit of the ones proposed in Engle and Marcucci, 2006 and presented in Section 3) tend to over-reject the null of no common factors. Moreover the approach becomes unfeasible when the number of series is larger than a few and when the multivariate ARCH also includes past volatilities. Consequently, usual canonical correlation should be avoided if the aim is to go above the investigation of commonalities in variances only between two series. Canonical correlations do not even work in some specific cases such as the BEKK (Baba *et al.*, 1995) model with reduced rank model with r factors as F-BEKK(p, q, r) (see Lin, 1992). In particular a

F-BEKK(0, 1, 1) reads as

$$\begin{aligned}\varepsilon_t &= H_t^{1/2} z_t, \\ H_t &= \Gamma_0 \Gamma_0' + \Gamma_1 \varepsilon_{t-1} \varepsilon_{t-1}' \Gamma_1', \\ &= \Gamma_0 \Gamma_0' + \gamma \varphi' \varepsilon_{t-1} \varepsilon_{t-1}' \varphi \gamma',\end{aligned}$$

where for $r = 1$ for instance we have $\text{Rank}(\Gamma_1) = 1$. For this setting a likelihood ratio test consisting to estimate the full and the restricted model is to be preferred. The Gaussian ML estimator is the one that maximizes the log likelihood function $L(\vartheta) = \sum_{i=1}^T l_t(\vartheta)$ with

$$l_t(\vartheta) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \ln |H_t(\vartheta)| - \frac{1}{2} \varepsilon_t' H_t^{-1}(\vartheta) \varepsilon_t.$$

For instance in a BEKK(0, 1), the unrestricted model is $H_t = \Gamma_0 \Gamma_0' + \Gamma_1 \varepsilon_{t-1} \varepsilon_{t-1}' \Gamma_1'$ and the restricted one $H_t = \Gamma_0 \Gamma_0' + \alpha \beta_1' \varepsilon_{t-1} \varepsilon_{t-1}' \beta_1 \alpha'$. Hecq *et al.* (2016) show that the LR test for the reduced rank BEKK with r factors, performs well. The advantage of this LR approach over the canonical correlation is that it can be easily generalized to F-BEKK(p, q, r). The main drawback of this approach might be the difficulty to estimate the model for large n .

3.3.2 Reduced-rank vector heterogeneous autoregressive models

Anderson and Vahid (2007) propose to look at information criteria for determining the presence and the number of principal component factors out of 21 Australian weekly stock return realized volatilities. When computed at daily frequency,⁸ the realized variances are obtained using $RV_t \equiv \sum_{i=1}^M r_{t,i}^2$, where $r_{t,i}$ are the high frequency intra-day returns, observed for M intra-day periods each day. For instance $M = 79$ for 5-min returns when the market is open between 9 a.m. to 4 p.m. The bipower variation $BV_t \equiv \frac{\pi}{2} \frac{M}{M-1} \sum_{i=2}^M |r_{t,i}| |r_{t,i-1}|$ is one of the measures of the integrated volatility that is designed to be robust to jumps. About a dozen of other specifications have been proposed in the literature. See, *i.a.*, Barndorff-Nielsen and Shephard (2004) and Bauwens, Hafner and Laurent (2012).

It turns out that working with realized volatility measures (that become new "observed" variables) is easier than estimating multivariate GARCH models. Moreover Gaussianity of errors is often assumed in the realized volatility literature. For daily volatility measures however the dynamic dependency is larger than a few lags. For instance, the univariate Heterogeneous Autoregressive model (HAR, see Corsi, 2009) captures the long range dependence observed in daily time series by a restricted autoregressive model of order 22.

In order to capture the dynamic interactions within a set of n daily realized volatility measures $Y_t^{(d)} \equiv (Y_{1,t}^{(d)}, \dots, Y_{n,t}^{(d)})'$, Bubák *et al.* (2011) and Souček and Todorova (2013) proposed, *inter alia*, to rely on a vector heterogeneous AR model (VHAR hereafter), an extension of the univariate HAR model. The vector $Y_t^{(d)}$ can include either the same kind of volatility measure (e.g., the realized variance) for different markets in a study of volatility co-movements or several volatility measures (realized variance, bipower variation, etc.) for the same market in order to construct an optimal linear combination like in Patton and Sheppard (2009) or Cubadda *et al.* (2017).

The Vector Heterogeneous Autoregressive model (VHAR) reads as follows:

$$Y_t^{(d)} = \beta_0 + \Phi^{(d)} Y_{t-1d}^{(d)} + \Phi^{(w)} Y_{t-1d}^{(w)} + \Phi^{(m)} Y_{t-1d}^{(m)} + \varepsilon_t, \quad t = 1, 2, \dots, T, \quad (18)$$

⁸ Anderson and Vahid (2007) use instead daily squared returns to estimate weekly realized volatilities.

where (d) , (w) , and (m) denote, respectively, time horizons of one day, one week (5 days a week), and one month (assuming 22 days within a month) such that

$$Y_t^{(w)} = \frac{1}{5} \sum_{j=0}^4 Y_{t-jd}^{(d)}, \quad Y_t^{(m)} = \frac{1}{22} \sum_{j=0}^{21} Y_{t-jd}^{(d)}.$$

Innovations ε_t are i.i.d. with $E(\varepsilon_t) = 0$, $E(\varepsilon_t \varepsilon_t') = \Sigma$ (positive definite), finite fourth moments.

One possible road to study reduced rank regression in a VHAR is to investigate the presence of common HAR features and search for a combination of $Y_t^{(d)}$ orthogonal to the past using instruments $Y_{t-1d}^{(d)}$, $Y_{t-1d}^{(w)}$ and $Y_{t-1d}^{(m)}$. This would be very similar to the investigation performed on multivariate GARCH but now we have a model for realized volatilities. In the presence of heteroskedasticity in the error terms of the VHAR, robust procedures presented in the previous section can be used. Under the SCCF restrictions, the reduced rank model that assumes a common left null space is

$$Y_t^{(d)} = \beta_0 + AC^{(d)}Y_{t-1d}^{(d)} + AC^{(w)}Y_{t-1d}^{(w)} + AC^{(m)}Y_{t-1d}^{(m)} + \varepsilon_t, \quad t = 1, 2, \dots, T, \quad (19)$$

such as the combination $A'_\perp Y_t^{(d)}$ does not have the HAR feature.

We can also look at commonalities in the right null spaces of the VHAR coefficient matrices by means of the VHAR-Index model (VHARI) of Cubadda *et al.* (2017). In such a case, assume that (18) can be rewritten as follows

$$Y_t^{(d)} = \beta_0 + \beta^{(d)}\omega'Y_{t-1d}^{(d)} + \beta^{(w)}\omega'Y_{t-1d}^{(w)} + \beta^{(m)}\omega'Y_{t-1d}^{(m)} + \varepsilon_t, \quad (20)$$

where ω is a $n \times r$ full-rank matrix. In terms of parameters, (20) needs $4(n \times r) - r^2$ instead of $n^2 \times 3$ in (18). Obviously, common volatility is allowed in the VHARI model if there exists a full-rank $n \times s$ (with $s < r$) matrix such that $\delta'\beta^{(i)} = 0$ for $i = d, w, m$.

Beyond the important aspect in terms of parsimony, there are two further motivations for using (20). First, the indexes $f_t^{(d)} = \omega'Y_{t-1d}^{(d)}$ obtained from (20) satisfy the property

$$f_t^{(w)} = \frac{1}{5} \sum_{j=0}^4 f_{t-jd}^{(d)}, \quad f_t^{(m)} = \frac{1}{22} \sum_{j=0}^{21} f_{t-jd}^{(d)}. \quad (21)$$

as for the observed univariate realized volatilities. Hence, the temporal cascade structure of the HAR model is preserved given that the weekly (monthly) index is equal to the weekly (monthly) moving average of the daily index. This would not be generally the case with traditional common feature models as in Engle and Marcucci (2006).

Second, premultiplying both side (20) by ω' yields

$$f_t^{(d)} = \omega'\beta_0 + \omega'\beta^{(d)}f_{t-1d}^{(d)} + \omega'\beta^{(w)}f_{t-1d}^{(w)} + \omega'\beta^{(m)}f_{t-1d}^{(m)} + \omega'\varepsilon_t, \quad (22)$$

which shows that the indexes themselves follow a VHAR model. When $r = 1$ the unique index is generated by an univariate HAR model. As we have already noticed in the previous section, this property is not shared by alternative methods to aggregate time series. Cubadda *et al.* (2017) offered a SA to estimate the parameters of the VHARI.

The VHARI model can be extended to include realized covariances in the analysis, as it necessary for the construction of optimal portfolios. A possible approach to follow is to apply the matrix log transformation to the realized covariance matrix as suggested by Bauer and Vorkink (2001).

4 Empirical illustrations

This section presents several empirical illustrations, focussing on guidelines for the reader.

4.1 Co-movements in housing prices

Based on the well known equivalence between the CCA and the full information maximum likelihood estimator (FIML) (see Hannan, 1967), we illustrate several ways to set up a system to be estimated by FIML. We consider the quarterly non seasonally adjusted house price index series released by the U.S. Federal Housing Finance Agency. Although the variables are available on a longer span, we study in this presentation the period 1990Q1 to 2020Q3. Indeed the 70s and the 80s data are more volatile; this impacts the robustness of the results, for instance by selecting a different VAR lag length for different samples. We study all transaction house prices series in California and Florida. Figure 1 shows the level of the series, investigations are performed for their logs.

We can observe a similar trending pattern in both prices. Surprisingly however, both usual ADF tests and HEGY seasonal unit root procedures on the logs of series suggest (5% level) that they are stationary around a deterministic trend; no deterministic dummies are relevant nor seasonal unit roots are present. Series being $I(0)$ around a deterministic time trend component, cointegration is not an issue for this investigation.

Figure 1: House price indexes

We consequently start with a bivariate VAR in log levels with an intercept and a deterministic linear trend. Information criteria favor $p = 4$ (with BIC) or 5 (AIC and HQ). The logs of the likelihood at its maximum for $p = 4$ and 5 for the unrestricted models are respectively -626.6425 and -619.7050 with $T = 123$ observations in both cases (we take the lags in the pre-sample period). Correlograms of residuals do not show any signs of serial correlation. Although it is not the issue of our empirical analyses, it is found that there is an unidirectional significant Granger causality relationship from California to Florida prices.

The first question we can ask ourself with this specification is the presence of a common deterministic components in the trend and the intercept. To do so we compare unrestricted systems

$$\begin{bmatrix} \ln Ca_t \\ \ln Fl_t \end{bmatrix} = \begin{bmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{bmatrix} \begin{bmatrix} 1 \\ t \end{bmatrix} + \sum_{i=1}^p \Phi_i \begin{bmatrix} \ln Ca_{t-i} \\ \ln Fl_{t-i} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

with restricted ones in which under RRR we have

$$\begin{bmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \begin{bmatrix} 1 & \beta_2 \end{bmatrix}.$$

There are two ways that are numerically equivalent in terms of their likelihood to evaluate a restricted system. We can write the pseudo causal representations that we estimate by FIML in

$$\begin{bmatrix} 1 & \gamma \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \ln Ca_t \\ \ln Fl_t \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ \tilde{\pi}_{21} & \tilde{\pi}_{22} \end{bmatrix} \begin{bmatrix} 1 \\ t \end{bmatrix} + \sum_{i=1}^p \tilde{\Phi}_i \begin{bmatrix} \ln Ca_{t-i} \\ \ln Fl_{t-i} \end{bmatrix} + \begin{bmatrix} \tilde{\varepsilon}_{1t} \\ \tilde{\varepsilon}_{2t} \end{bmatrix}$$

which means to add the contemporaneous variable $\ln Fl_t$ in the right hand side of the first equation and delete the intercept and the trend. Alternatively one can also formalize a factor like structure

$$\begin{bmatrix} \ln Ca_t \\ \ln Fl_t \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \begin{bmatrix} 1 & \beta_2 \end{bmatrix} \begin{bmatrix} 1 \\ t \end{bmatrix} + \sum_{i=1}^p \tilde{\Phi}_i \begin{bmatrix} \ln Ca_{t-i} \\ \ln Fl_{t-i} \end{bmatrix} + \begin{bmatrix} \tilde{\varepsilon}_{1t} \\ \tilde{\varepsilon}_{2t} \end{bmatrix}.$$

Both approaches provide the same value of the likelihood. The advantage of the first one is that we can also obtain standard errors of the combination of the series, namely on $\hat{\gamma}$. The restricted likelihoods are -627.8792 and -621.4003 for $p = 4$ and $p = 5$. LR tests are 2.473 and 3.39 for $p = 4$ and $p = 5$. They follow a $\chi^2_{(1)}$, hence we do not reject the null of common deterministic components in both cases at a 5% significance level.

Let us now leave the deterministic terms unrestricted to focus on the short run dynamics. We do not impose a common left null space to both dynamics and deterministic coefficient matrices. The following table provides the values of the likelihood under SCCF as well as for several non synchronous models. It emerges that every restriction is rejected at a 5% significance level but the PSCCF(4) in a VAR(5). This reflects our hesitation to choose the lag length. With $p = 5$, there is a trivial common feature vector annihilating the last coefficient matrix.

Table 1 - Common Features tests statistics

	SCCF	PSCCF(1)	PSCCF(2)	PSCCF(3)	PSCCF(4)	Unrestricted
$p = 4$	-877.3379	-654.2018	-633.4261	-630.7047	—	-626.6425
#restrictions	7	5	3	1		
$p = 5$	-875.9611	-652.8442	-630.1269	-627.4759	-621.0366	-619.7050
#restrictions	9	7	5	3	1	

From the previous table we conclude that there aren't any common cyclical features generating the dynamics of the system.

Similarly we can also look at multivariate index models that seeks for a common right null space in the

matrix lag coefficients. The system with $p = 4$ and no restrictions on the deterministic components is

$$\begin{aligned} \begin{bmatrix} \ln Ca_t \\ \ln Fl_t \end{bmatrix} &= \begin{bmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{bmatrix} \begin{bmatrix} 1 \\ t \end{bmatrix} + \\ &\begin{bmatrix} a_{11} \\ \alpha_{21} \end{bmatrix} \begin{bmatrix} 1 & c \end{bmatrix} \begin{bmatrix} \ln Ca_{t-1} \\ \ln Fl_{t-1} \end{bmatrix} + \\ &\dots\dots\dots \\ &+ \begin{bmatrix} a_{14} \\ \alpha_{24} \end{bmatrix} \begin{bmatrix} 1 & c \end{bmatrix} \begin{bmatrix} \ln Ca_{t-4} \\ \ln Fl_{t-4} \end{bmatrix} \\ &+ \begin{bmatrix} \tilde{\varepsilon}_{1t} \\ \tilde{\varepsilon}_{2t} \end{bmatrix}. \end{aligned}$$

We obtain logs of the likelihoods that are respectively equal to -887.9438 with $p = 4$ and -880.3337 with $p = 5$. The MAI is consequently rejected by the data as well. In this application the only source of commonalities apparently comes from the deterministic components.

4.2 Common cycles between Chilean and Peruvian economic growths

We investigate the degree of long and short-run dependence among two Latin American economies, Peru and Chile. In particular, we consider annual real gross domestic products data (at 2011 prices) from 1900 to 2018 downloaded from the Total Economy Database⁹. Figures 2 and 3 display the log-levels and the growth rates of the series. We stick to the bivariate case for the sake of simplicity and we do not aim to perform a formal study on the convergence mechanisms linking Latin American countries (see for that e.g. Hecq (2002) or Christensen et al. (2011)).

Figure 2: Log levels of GDPs

The log-levels of the variables are clearly non stationary and one does not reject the null of a unit root in both log-level series. p -values are respectively of 0.34 and 0.35 for Chile and Peru using ADF tests with an intercept and a linear trend. A VAR(3) characterizes parsimoniously the log-levels of the bivariate process according to information criteria. Using the Johansen's trace test we do not reject the null of no

⁹University of Groningen www.eco.rug.nl/ggdc.

Figure 3: Growth rates of GDPs

cointegration, the p -value associated with the null hypothesis of no cointegration is 0.53. From Figure 2, it might be expected that these series do not share a common tendency.

Consequently we proceed with a VAR(2) for the growth rates of the GDPs. We report next the values of the unrestricted coefficient matrices, with t -tests underneath. In the event that there would exist a SCCF vector, a quick look at those coefficient matrices reveals that the vector should not be so different from $[1, -1]'$. This can be seen in the estimated parameters where the second row is roughly equal to the first one. More formally, the likelihood ratio test statistic for a reduced rank is 0.074 which is much lower than the $\chi^2_{(3)}$ critical values at any usual significance levels. The common feature relationship with t -test in parentheses is $\Delta \ln Chl_t - 0.9 \Delta \ln Per_t$, the RRR has 3 parameters less than the unrestricted VAR.

$$\begin{bmatrix} \Delta \ln Chl_t \\ \Delta \ln Per_t \end{bmatrix} = \begin{bmatrix} 0.003 & 0.402 \\ (0.035) & (2.416) \\ 0.019 & 0.441 \\ (0.345) & (4.680) \end{bmatrix} \begin{bmatrix} \Delta \ln Chl_{t-1} \\ \Delta \ln Per_t \end{bmatrix} + \begin{bmatrix} -0.144 & -0.098 \\ (-1.488) & (-0.585) \\ -0.156 & -0.076 \\ (-2.850) & (-0.809) \end{bmatrix} \begin{bmatrix} \Delta \ln Chl_{t-2} \\ \Delta \ln Per_{t-2} \end{bmatrix} + \begin{bmatrix} 0.011 \\ (1.500) \\ 0.014 \\ (3.354) \end{bmatrix} + \begin{bmatrix} \hat{\varepsilon}_{1t} \\ \hat{\varepsilon}_{2t} \end{bmatrix}$$

5 Conclusion

The literature on the use of RRR models in economics and finance is vast. We have reviewed in this chapter some of the most important specifications for the analysis of stationary or stationarized multivariate time series. Unfortunately, we were not able to cover numerous contributions, both of theoretical and empirical nature, that have been published in many good journals over time. Indeed, we focused on the specifications that we believe are the most fruitful in time-series econometrics: common features, multivariate index models, and common heteroskedasticity.

The take away message is that RRR is a very useful tool to detect comovements (not only the long-run ones), to obtain efficiency gains through the imposition of parameter restrictions, and to boost forecast

accuracy. The drawback, but we personally are not sure that it is a real one, is that reduced rank restrictions might not always be present in the data at hand, making the use of RRR in empirical research more challenging with respect to tools that impose restrictions to the data without a formal pre-testing.

References

- [1] AHN, S.K. (1997), Inference of Vector Autoregressive Models with Cointegration and Scalar Components, *Journal of the American Statistical Association*, 92, 350-356.
- [2] AHN, S.K., AND G.C. REINSEL (1988), Nested Reduced-rank Autoregressive Models for Multiple Time-series, *Journal of the American Statistical Association*, 83, 849-856.
- [3] ANDERSON, T.W. (1951), Estimating Linear Restrictions on Regression Coefficients for Multivariate Normal Distributions. *Annals of Mathematical Statistics*, 22, 327-351.
- [4] ANDERSON, T.W. (1984), *An Introduction to Multivariate Statistical Analysis*, Second Edition. New York: Wiley.
- [5] ANDERSON, T.W. (1999), Asymptotic Theory for Canonical Correlation Analysis, *Journal of Multivariate Analysis*, 70, 1-29.
- [6] ANDERSON, T.W. (2002), Canonical Correlation Analysis and Reduced Rank Regression in Autoregressive Models, *Annals of Statistics*, 30, 1134-1154.
- [7] BABA, Y., R.F. ENGLE, D.F. KRAFT, AND K.F. KRONER (1989), Multivariate Simultaneous Generalized ARCH, DP 89-57 San Diego.
- [8] BAUER, G.H., AND K. VORKINK (2011), Forecasting multivariate realized stock market volatility, *Journal of Econometrics*, 160, 93-101
- [9] BAUWENS, L., S. LAURENT AND J.V.K. ROMBOUTS (2006), Multivariate GARCH Models: a Survey, *Journal of Applied Econometrics*, 21, 79-109.
- [10] BAUWENS, L., HAFNER, CH. AND S. LAURENT (2012), Volatility Models, *Handbook of Volatility Models and Their Applications* (eds. Bauwens, L., Hafner, Ch. and S. Laurent), Wiley, Ch. 1
- [11] BEINE, M. AND A. HECQ (1999), Inference in Codependence: Some Monte Carlo Results and Applications, *Annales d'Economie et de Statistique*, 54, 69-90.
- [12] BEINE, M., CANDELON, B. AND A. HECQ (2000), Determining A Perfect European Optimum Currency Area Using Common Cycle, *Empirica*, 27, 115-132.
- [13] BARNDORFF-NIELSEN, O.E. AND N. SHEPHARD, ECONOMETRIC (2004), Analysis of Realized Covariation: High Frequency Based Covariance, Regression, and Correlation in Financial Economics, *Econometrica*, 72, 885-925.
- [14] BERNARDINI E. AND G. CUBADDA (2015), Macroeconomic Forecasting and Structural Analysis through Regularized Reduced-Rank Regression, *International Journal of Forecasting*, 31, 682-691.

- [15] BIERENS, H. (2000), Nonparametric Nonlinear Cotrending Analysis, with an Application to Interest and Inflation in the United States, *Journal of Business & Economic Statistics*, 18, 323–337.
- [16] BOSWIJK, H.P., (1995), Identifiability of cointegrated systems, *Tinbergen Institute Working Paper*, 95/78.
- [17] BOSWIJK, H.P, AND J.A. DOORNIK (2004), Identifying, estimating and testing restricted cointegrated systems: An overview, *Statistica Neerlandica*, 58, 440–465.
- [18] BUBAK, V., KOVCENDA, E. AND F. ZIKES (2011), Volatility transmission in emerging European exchange markets, *Journal of Banking & Finance*, 35, 2829–2841.
- [19] CANDELON, B., HECQ, A. AND W. VERSCHOOR (2005), Measuring common cyclical features during financial turmoil: Evidence of interdependence not contagion, *Journal of International Money and Finance*, 24, 1317–1334.
- [20] CARRIERO, A. KAPETANIOS, G., AND M. MARCELLINO (2011), Forecasting Large Datasets with Bayesian Reduced Rank Multivariate Models, *Journal of Applied Econometrics*, 26, 735–761.
- [21] CARRIERO, A. KAPETANIOS, G., AND M. MARCELLINO (2016), Structural analysis with Multivariate Autoregressive Index models, *Journal of Econometrics*, 192, 332–348.
- [22] CENTONI, M., AND G. CUBADDA (2003), Measuring the business cycle effects of permanent and transitory shocks in cointegrated time series. *Economics Letters* 80, 45–51.
- [23] CENTONI M., AND G. CUBADDA (2015), Common feature analysis of economic time series: An overview and recent developments, *Communications for Statistical Applications and Methods*, 22, 1–20.
- [24] CENTONI M., CUBADDA G., AND A. HECQ (2007), Common shocks, common dynamics, and the international business cycle, *Economic Modelling*, 24, 149–166.
- [25] CHANG, J., GUO, B., AND Q YAO (2015), High dimensional stochastic regression with latent factors, endogeneity and nonlinearity, *Journal of Econometrics*, 189, 297–312.
- [26] CHEVILLON, G., HECQ, A. AND S. LAURENT (2018), Generating Univariate Fractional Integration within a Large VAR(1)”, *Journal of Econometrics*, 204(1),
- [27] CHRISTENSEN, T., HURN, S. AND A. PAGAN (2011), Detecting Common Dynamics in Transitory Components, *Journal of Time Series Econometrics*, 2011, vol. 3, issue 1.
- [28] CORSI, F. (2009), A Simple Approximate Long-Memory Model of Realized Volatility, *Journal of Financial Econometrics*, 7, 174–196.
- [29] CUBADDA, G. (1999), Common Cycles in Seasonal Non-stationary Time Series, *Journal of Applied Econometrics*, 14, 273–291.
- [30] CUBADDA, G. (2001), Common Features In Time Series With Both Deterministic And Stochastic Seasonality, *Econometric Reviews*, 20, 201–216.
- [31] CUBADDA, G. (2007), A Unifying Framework for Analyzing Common Cyclical Features in Cointegrated Time Series, *Computational Statistics and Data Analysis*, 52, 896–906.

- [32] CUBADDA, G., AND B. GUARDABASCIO (2019), Representation, Estimation and Forecasting of the Multivariate Index-Augmented Autoregressive Model, *International Journal of Forecasting*, 35, 67–79.
- [33] CUBADDA G., GUARDABASCIO B., AND A. HECQ (2017), A Vector Heterogeneous Autoregressive Index Model for Realized Volatility Measures, *International Journal of Forecasting*, 33, 337–344.
- [34] CUBADDA, G. AND A. HECQ (2001), On Non-Contemporaneous Short-run Co-movements, *Economics Letters*, 73, 389–397.
- [35] CUBADDA, G. AND A. HECQ (2011), Testing for Common Autocorrelation in Data Rich Environments, *Journal of Forecasting*, 30, 325–335.
- [36] CUBADDA, G. AND A. HECQ (2021), Dimension Reduction for High Dimensional Vector Autoregressive Models, *arXiv.org*, arXiv:2009.03361v2.
- [37] CUBADDA, G., HECQ, A. AND F.C. PALM (2008), Macro-panels and Reality, *Economics Letters*, 99, 537–540.
- [38] CUBADDA, G., HECQ A. AND F.C. PALM (2009), Studying Interactions without Multivariate Modeling, *Journal of Econometrics*, 148, 25–35.
- [39] CUBADDA G., HECQ A., AND S. TELG (2019), Detecting Co-Movements in Noncausal Time Series, *Oxford Bulletin of Economics and Statistics*, 81, 697–715.
- [40] DIEBOLD, F.X. (1986), Testing for Serial Correlation in the Presence of ARCH, *Proceedings of the ASA Business and Economic Statistics Section*, 323–328.
- [41] ENGLE, R.F. AND R. SUSMEL (1993), Common Volatility in International Equity Markets, *Journal of Business and Economic Statistics*, 11, 167–175.
- [42] ENGLE, R. F. (2002), Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models, *Journal of Business and Economic Statistics* 20, 339–350.
- [43] ENGLE, R. F. AND J. MARCUCCI (2006), A Long-run Pure Variance Common Features Model for the Common Volatilities of the Dow Jones, *Journal of Econometrics* 132, 7–42.
- [44] ENGLE, R.F., AND S. KOZICKI (1993), Testing for Common Features (with comments), *Journal of Business and Economic Statistics*, 11, 369–395.
- [45] ENGLE, R.F., AND C.W.J. GRANGER (1987), Co-integration and Error correction: Representation, Estimation and Testing, *Econometrica*, 55, 251–276.
- [46] ENGLE, R.F. AND HYLLEBERG, S. (1996), Common Seasonal Features: Global Unemployment, *Oxford Bulletin of Economics and Statistics*, 58, 615–630.
- [47] FRANCHI, M. , AND P. PARUOLO (2011), A Characterization of Vector Autoregressive Processes with Common Cyclical Features, *Journal of Econometrics*, 163, 105–117.
- [48] GOETZ, T., HECQ, A. AND J.-P. URBAIN (2013), Testing for Common Cycles in Non-Stationary VARs with Varied Frequency data, *Advances in Econometrics*, 32, 361–393

- [49] GOETZ, T., HECQ, A. AND S. SMEEKES (2016), Testing for Granger-causality in large mixed-frequency VARs, *Journal of Econometrics*, 193, (2), 418-432.
- [50] GOURIÉROUX, CH. AND I. PEAUCELLE (1993), *Séries codépendantes : application à l'hypothèse de parité du pouvoir d'achat*, in *Macroéconomie, développements récents*, éd. Economica, 285-306.
- [51] GUILLÉN, O., HECQ, A., ISSLER, J. AND D. SARAIVA (2015), Forecasting Multivariate Time Series under Present-Value-Model Short- and Long-run Co-movement Restrictions, *International Journal of Forecasting*, 31, (3), 862-875.
- [52] HANNAN, E.J. (1967), Canonical Correlation and Multiple Equation Systems in Economics, *Econometrica*, 123-138.
- [53] HECQ, A. (1996), IGARCH effect on Autoregressive Lag Length Selection and Causality Tests, *Applied Economics Letters*, 3, 317-323.
- [54] HECQ, A. (2000), *Common Cyclical Features in Multiple Time Series and Panel Data. Methodological Aspects and Applications*, Ph.D. thesis.
- [55] HECQ, A. (2002), Common Cycles and Common Trends in Latin America, *Medium Econometrische Toepassingen*, 10, 20-25.
- [56] HECQ, A. AND J. ISSLER (2012), A Common-feature approach for testing present-value restrictions with financial data, *FGV EPGE Economics Working Papers*, No 728.
- [57] HECQ, A., LAURENT, S. AND F. PALM (2012), Common Intraday Periodicity, *Journal of Financial Econometrics*, 10, 325-353.
- [58] HECQ, A., LAURENT, S. AND F. PALM (2016), On the Univariate Representation of the BEKK Model with Common Factors, *Journal of Time Series Econometrics*, 8, 91-113.
- [59] HECQ, A., F. PALM AND J.P. URBAIN (2000), Comovements in International Stock Markets: What can we Learn from a Common Trend-Common Cycle Analysis? , *De Economist*, 148, 395-406.
- [60] HECQ, A., PALM, F., AND J. URBAIN (2006), Common Cyclical Features Analysis in VAR Models with Cointegration, *Journal of Econometrics*, 132, 117-141.
- [61] HENDRY, D.F. (1999), Co-Breaking, in *Forecasting Non-stationary Economic Time Series*, eds. D.F. Hendry and M.P. Clements, MIT Press, Cambridge MA.
- [62] IZENMAN, A. (2008), *Modern Multivariate. Statistical Techniques: Regression, Classification and Manifold Learning*, Springer, New York.
- [63] JOHANSEN, S. (1996), *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, Second Edition.
- [64] KADIYALA, K.R. AND S. KARLSSON (1997), Numerical methods for estimation and inference in Bayesian VAR-models, *Journal of Applied Econometrics*, 12 99-132.
- [65] KLEIBERGEN, F. AND R. PAAP (2006), Generalized reduced rank tests using the singular value decomposition, *Journal of Econometrics*, 133, 97-126.

- [66] LAM, C. AND Q. YAO (2012), Factor modeling for high-dimensional time series: Inference for the number of factors, *Annals of Statistics*, 40, 694-726.
- [67] LAM, C., YAO, Q., AND N. BATHIA, (2011), Estimation of latent factors for high-dimensional time series, *Biometrika*, 98 901-918.
- [68] LIPPI, M. (2019), Time-domain approach in high-dimensional dynamic factor models, in *Oxford research encyclopedia of economics and finance*, Oxford University Press. New York.
- [69] LUCAS, R. (1977), Understanding Business Cycles. *Carnegie-Rochester Conference Series on Public Policy*, 5: 7-29.
- [70] OBERHOFER, W. AND J. KMENTA (1974), A general procedure for obtaining maximum likelihood estimates in generalized regression models, *Econometrica*, 42, 579-590.
- [71] PARUOLO, P. (2006), Common trends and cycles in I(2) VAR systems, *Journal of Econometrics*, 132, 143-168.
- [72] RAVIKUMAR, B., RAY, S. AND E. SAVING (2000), Robust Wald tests in Sur systems with adding-up restrictions, *Econometrica*, 68, 715-720.
- [73] REINSEL, G. (1983), Some Results on Multivariate Autoregressive Index Models, *Biometrika*, 70, 145-15.
- [74] REINSEL, G., AND S.K AHN (1992), Vector Autoregressive Models with Unit Roots and Reduced Rank Structure: Estimation. Likelihood Ratio Test, and Forecasting, *Journal of Time Series Analysis*, 13, 353-375
- [75] REINSEL, G., AND R. VELU (1998), *Multivariate Reduced-Rank Regression: Theory and Applications*, Springer, New York.
- [76] SCHLEICHER, C. (2007), Codependence in Cointegrated Autoregressive Models, *Journal of Applied Econometrics*, 22, 137-159.
- [77] SOUCEK, M. AND N. TODOROVA (2013), Realized volatilities transmission between crude oil and equity futures markets: A multivariate HAR approach, *Energy Economics*, 40, 586-597.
- [78] TIAO, G. C. AND R. S. TSAY (1989), Model Specification in Multivariate Time Series (with comments), *Journal of Royal Statistical Society, Series B*, 51, 157-213.
- [79] VAHID, F., AND R.F. ENGLE (1993), Common Trends and Common Cycles, *Journal of Applied Econometrics*, 8, 341-360.
- [80] VAHID, F., AND R.F. ENGLE (1997), Codependent Cycles, *Journal of Econometrics*, 80, 199-221.
- [81] VAHID, F., AND J.V. ISSLER (2002), The Importance of Common Cyclical Features in VAR Analysis: a Monte-Carlo Study, *Journal of Econometrics*, 109, 341-363.

RECENT PUBLICATIONS BY *CEIS Tor Vergata*

Herding and Anti-Herding Across ESG Funds

Rocco Ciciretti, Ambrogio D'Alò and Giovanni Ferri

CEIS Research Paper, 524, November 2021

Experimental Analysis of Endogenous Institutional Choice: Constantly Revealing versus Ad-hoc Contracting

Daniela Di Cagno, Lorenzo Ferrari, Werner Güth and Vittorio LaroCCA

CEIS Research Paper, 523, November 2021

Institutions and Economic Development: New Measurements and Evidence

Leonardo Becchetti and Gianluigi Conzo

CEIS Research Paper, 522, November 2021

Institutions and Economic Development: New Measurements and Evidence

Esther Acquah, Lorenzo Carbonari, Alessio Farcomeni and Giovanni Trovato

CEIS Research Paper, 521, November 2021

The Transmission Mechanism of Quantitative Easing: A Markov-Switching FAVAR Approach

Luisa Corrado, Stefano Grassi and Enrico Minnella

CEIS Research Paper, 520, October 2021

Keeping the Agents in the Dark: Private Disclosures in Competing Mechanisms

Andrea Attar, Eloisa Campioni, Thomas Mariotti and Alessandro Pavan

CEIS Research Paper, 519, October 2021

Modelling Cycles in Climate Series: the Fractional Sinusoidal Waveform Process

Tommaso Proietti and Federico Maddanu

CEIS Research Paper, 518, October 2021

Modelling and Estimating Large Macroeconomic Shocks During the Pandemic

Luisa Corrado, Stefano Grassi and Aldo Paolillo

CEIS Research Paper, 517, October 2021

Three Liquid Assets

Nicola Amendola, Lorenzo Carbonari and Leo Ferraris

CEIS Research Paper, 516, October 2021

Efficient Nonparametric Estimation of Generalized Autocovariances

Alessandra Luati, Francesca Papagni and Tommaso Proietti

CEIS Research Paper, 515, October 2021

DISTRIBUTION

Our publications are available online at www.ceistorvergata.it

DISCLAIMER

The opinions expressed in these publications are the authors' alone and therefore do not necessarily reflect the opinions of the supporters, staff, or boards of CEIS Tor Vergata.

COPYRIGHT

Copyright © 2021 by authors. All rights reserved. No part of this publication may be reproduced in any manner whatsoever without written permission except in the case of brief passages quoted in critical articles and reviews.

MEDIA INQUIRIES AND INFORMATION

For media inquiries, please contact Barbara Piazzzi at +39 06 72595652/01 or by e-mail at piazzzi@ceis.uniroma2.it. Our web site, www.ceistorvergata.it, contains more information about Center's events, publications, and staff.

DEVELOPMENT AND SUPPORT

For information about contributing to CEIS Tor Vergata, please contact at +39 06 72595601 or by e-mail at sagr.ceis@economia.uniroma2.it