Common Features in Economics and Finance: An Overview of Recent Developments

Giovanni Urga

CEA@Cass Working Paper Series
WP–CEA–07-2006
Common Features in Economics and Finance: An Overview of Recent Developments

Giovanni Urga
Faculty of Finance, Centre for Econometric Analysis, Cass Business School, 106 Bunhill Row, London EC1Y 8TZ, U.K.
(g.urga@city.ac.uk)
October 20, 2006

Abstract

This introductory paper offers an overview of some developments in the common features literature since the publication of the seminal paper by Engle and Kozicki in the Journal of Business & Economic Statistics in 1993, with the aim of highlighting the unifying theme of the contributions in this volume.
1 INTRODUCTION

Economic or financial time series may exhibit many distinctive features such as trends, cycles, serial correlation, seasonality, time varying volatility, breaks and non-linearities. The presence of the same type of feature in a group of variables conveys valuable information that is useful for developing parsimonious models and provides evidence regarding the relevance of economic theories which imply such common features. A typical example is when two series each have a stochastic trend. If these trends are shared, that is when the series are cointegrated, we conclude that there exists long run relationship between the variables, and then the information in both series can be pooled to produce a more accurate study of their common stochastic trends.

The notion of the common feature was formalized by Engle and Kozicki (1993, EK henceforth). If a group of series possesses a feature, this feature is said to be common if a linear combination of the series does not have the feature. In the example of cointegration (Engle and Granger 1987; Johansen 1988), all variables have stochastic trends but there are some linear combinations of the variables that do not have stochastic trends. Other examples include common serial correlation (EK) when all variables are serially correlated but some linear combinations of them are white noise, common cycles (Vahid and Engle 1993) when deviations from trend in all variables are cyclical but some linear combinations have no cycles, codependence (Gouriéroux and Peaucelle, 1988; Tiao and Tsay, 1989; Vahid and Engle 1997) when some linear combinations of variables have memory shorter than the memory of any of the constituent series, common seasonality (Engle and Hylleberg, 1996) when each series is seasonal but there are some combinations of them that are non-seasonal. Further developments include common structural breaks (Hendry 1996), common non-linearity (Anderson and Vahid 1998), common jumps (Barndorff-Nielsen and Shephard 2004), common time varying volatility (Engle and Marcucci 2006) and weak forms of common serial correlation features as defined in Cubadda and Hecq (2001) and Hecq, Palm and Urbain (2006).

Let us consider the \( n \)-dimensional time series processes \( y_t \) with \( t = 1, \ldots, T \). EK define a feature when the following three axioms are satisfied:

**Axiom 1** If each of the \( y_t \) has (does not have) the feature, then each \( \lambda y_t \) will have (will not have) the feature for any \( \lambda \neq 0 \).
Axiom 2 If \( y_{it} \) does not have the feature and \( y_{jt} \) does not have the feature, then \( z_t = y_{it} + y_{jt} \) will not have the feature; \( \forall i \neq j, i, j = 1, \ldots, n \).

Axiom 3 If \( y_{it} \) does not have the feature but \( y_{jt} \) does have the feature, then \( z_t = y_{it} + y_{jt} \) will have the feature.

EK provide the following definition of a common feature:

Definition 1 A feature present in each of a group of series is said to be common to those series if there exists a nonzero linear combination of the series that does not have the feature. Such a linear combination is called a "cofeature" combination and the vector which represents it is called a cofeature vector. For an \( n \)-dimensional system \( y_t = (y_{1t}, \ldots, y_{nt})' \) measured for \( t = 1, \ldots, T \), there can exist \( s < n \) linearly independent cofeature vectors. The collection of all linearly independent cofeature vectors form the \( n \times s \) matrix \( \delta \) where \( \{\delta' y_t\}_{t=1}^T \) does not have the feature. The range of the matrix \( \delta \) which defines the subspace of all possible cofeature vectors is called the cofeature space.

This definition implies a common factor structure for the \( n \)-dimensional series \( y_t = (y_{1t}, \ldots, y_{nt})' \). In the case that the feature relates to the conditional mean of \( y_t \), we have

\[
y_t \begin{bmatrix} (n \times 1) \\ (n \times (n-s)) \\ (n-s) \times 1 \\ (n \times 1) \end{bmatrix} = B \begin{bmatrix} F_t \\ u_t \end{bmatrix}, \quad t = 1 \ldots T, \tag{1}
\]

where the common factors \( F_t \) have a feature and \( u_t \) does not have the feature.

For example, when variables are cointegrated, \( F_t \) are common I(1) trends and \( u_t \) are I(0). In the analysis of common serial correlation, \( F_t \) are serially correlated and \( u_t \) are white noise. In the case of common seasonality, \( F_t \) are seasonal and \( u_t \) non-seasonal. The factor loading matrix \( B \) captures the influence of common factors on each variable, and \( s \) linearly independent vectors \( \delta \) such that \( \delta' B = 0 \) span the "cofeature space". The latter is the space of all vectors such that \( \delta' y_t \) is not influenced by \( F_t \) and hence does not exhibit the common feature of elements of \( y_t \).

This special issue of the *Journal of Business & Economic Statistics* contains a collection of papers from the conference on "Common Features in London", held at the Cass Business School on 16-17 December 2004. The conference was the third opus of a series of meetings, with the first one held.
in Rio in 2002 and the second in Maastricht in 2003. This introductory paper is not meant to be a comprehensive survey of the common features literature. Rather, it aims to provide an overview of historical developments to offer readers a background to follow the contributions in this volume. We feel this special issue, together with the 2006 special issue of the *Journal of Econometrics* on "Common Features", edited by Anderson, Issler and Vahid (2006), reflects important developments in the topic since the original contributions in the *Journal of Business & Economic Statistics* issue in 1993, and illustrate how many exciting issues remain to be developed in future research. Other recent contributions such as Brooks (2006), Bauwens, Laurent and Rombouts (2006), Johansen (2006a), Vahid (2006) nicely complement this special issue.

The paper is organized as follows. Section 2 provides an overview of the common features literature, including a summary of the papers of the special issue. Section 3 concludes and offers some thoughts on possible future development of the topic.

## 2 COMMON FEATURES

### 2.1 Common Trends and Common Cycles

Two important features of macroeconomic time series are their trends and cycles. An early approach to joint analysis of common trends and cycles as common features was proposed by Vahid and Engle (1993), who derive the restrictions that such common features impose on the reduced form VAR dynamics of the system, and design tests for such common features and suggest procedures to recover trends and cycles from the restricted reduced form.

When two or more series share a trend, the series are cointegrated and thus there exists a long run relationship between these variables. A simple VAR representation of order $p$ for an $n$-dimensional vector of I(1) time series in $y_t$ is

$$A(L)y_t = \varepsilon_t, \quad t = 1, ..., T,$$

(2)

where $A(L) = I_n - \sum_{i=1}^{p} A_i L^i$, $\varepsilon_t \sim NID(0, \Omega)$ (Johansen 1988). Let us assume that the components of the vector $y_t$ are cointegrated of order 1, 1, [denoted $y_t \sim CI(1, 1)$] with cointegrating rank equal to $r$. Thus, there exists
a vector equilibrium-correction model (VECM),

\[ \Delta y_t = \alpha \delta'_\beta y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t, \]

where \( \alpha \) and \( \delta_\beta \) are \((n \times r)\) matrices with rank equal to \( r \) such that \( \alpha \delta'_\beta = -A(1), \Gamma(L) = I_n - \sum_{i=1}^{p-1} \Gamma_i L^i, \) and \( \Gamma_i = -\sum_{j=i+1}^p A_j \) for \( i = 1, 2, \ldots, p - 1. \) The columns of \( \delta_\beta \) span the cointegrating space and the elements of \( \alpha \) are the corresponding adjustment coefficients. Obviously, cointegration is a leading example of common features, a framework in which there exist linear combinations of non-stationary variables (the stochastic trends feature) which are instead stationary.

There are plenty of contributions on cointegration at zero frequency, seasonal cointegration, periodic cointegration and fractional cointegration. Interested readers can also refer to the overviews and books on the topic by Johansen (1995, 2006a), Hylleberg (1992), Franses (1996), Robinson (2003), and Banerjee and Urga (2005).

Finally, the following contributions are worth mentioning. Chapman and Ogaki (1993) deal with co-trending regressions, Bierens (2000) with nonparametric non-linear co-trending rank, and Hatanaka and Yamada (2003) with co-trending for segmented trends.

Beyond and/or in addition to the presence of long-run co-movements, cyclical swings appear to be common over a set of variables. Gouriéroux and Peaucelle (1988), Tiao and Tsay (1989) and EK have analyzed the existence of short-run co-movements between stationary time series or between first differences of cointegrated I(1) series. EK developed the concept of Serial Correlation Common Features (SCCF). [Developments in the statistical literature on reduced rank regression that preceded the EK work on serial correlation common features are in Pena and Box (1987), Ahn and Reinsel (1988), Tsay (1989), Tsay and Tiao (1985), and Velu, Reinsel and Wichern (1986).] SCCF arise when there exists a cofeature matrix \( \delta_s \) such that \( \delta'_s \Delta y_t \) is a \( s \)-dimensional innovation process and that consequently the following restrictions to the VECM in (3) jointly hold: \( \delta'_s \alpha = 0 \) and \( \delta'_s \Gamma_i = 0, \) \( i = 1, \ldots, p - 1. \) Following Hecq (2006), the presence of SCCF allows us to rewrite the VECM into the common factor representation

\[ \Delta y_t = \delta_{s} A' W_t + \varepsilon_t = \delta_{s} F_t + \varepsilon_t, \]
where \( \delta_{ls} \) is a full-rank \( n \times (n - s) \) matrix such that \( \delta'_{ls} \delta_{ls} = 0 \) and \( W_t = (y'_{t-1}\beta, \Delta y'_{t-1}, \ldots, \Delta y'_{t-p+1})' \). Formulation (4) highlights the fact that SCCF implies that the system has a reduced number of propagation mechanisms for the transmission of the information contained in the past.

It is worth noting that the conditions underlying SCCF models are often violated due to, for instance, temporal aggregation (as in Marcellino 1999), seasonal adjustment (Cubadda 1999; Hecq 1998), GARCH processes (Candelon, Hecq and Verschoor 2005), and linear and log-linear transformations (Corradi and Swanson 2006).

The need for less stringent restrictions to allow, amongst others, for adjustment delays resulting from the presence of adjustment costs and habit formation (Ericsson 1993) motivated the first variant of the SCCF, namely the Codependent Cycle Feature (CCF) proposed by Gouriéroux and Peaucelle (1992), Kugler and Neusser (1993), recently reconsidered by Schleicher (2006). In the case of an adjustment delay of one period, i.e. codependence of order one, there exist combinations of \( \Delta y_t \) that are innovations with respect to the information set available at time \( t - 2 \) and not \( t - 1 \) as in the SCCF case. Following Hecq (2006), if there exists a codependence matrix \( \delta'_c \) such that \( \delta'_c \Delta y_t \) is a \( VMA(1) \), the common factor representation for the CCF of order one is

\[
\Delta y_t = \delta_{lc} \tilde{A}' W_{t-1} + \varepsilon_t + C_1 \varepsilon_{t-1},
\]  

(5)

where \( \delta_{lc} \) is a full-rank \( n \times (n - s) \) matrix such that \( \delta'_c \delta_{lc} = 0 \), \( W_{t-1} = (y'_{t-2}\beta, \Delta y'_{t-2}, \ldots, \Delta y'_{t-p})' \).

Another approach similar to CCF is the Polynomial Serial Correlation Common Features (PSCCF) proposed by Cubadda and Hecq (2001). This approach models the situation in which we have no SCCF because the short run co-movements are not contemporaneous, but there may exist linear combinations of \( \Delta y_t \) and its lags that are innovations with respect to the past.

Finally, Hecq, Palm and Urbain (2000a,b, 2002, 2006) propose the so-called Weak Form (WF) reduced rank structure common cycle feature, which arises when \( s \) linear combinations of \( \Delta y_t \) in deviation from the equilibrium-correction terms are white noise. The restriction involved is \( \delta'_{w} \Gamma_i = 0, i = 1 \ldots p - 1 \) and the common factor representation can be written as

\[
\Delta y_t = \alpha \delta'_{\beta} y_{t-1} + \delta_{lw} \tilde{A}' W_{t-1} + \varepsilon_t,
\]  

(6)

where \( \delta_{lw} \) is a full-rank \( n \times (n - s) \) matrix and \( \tilde{W}_{t-1} = (\Delta y'_{t-1}, \ldots, \Delta y'_{t-p+1})' \).
The choice amongst models depends on restrictions satisfied by the data. Many testing procedures that have been proposed are based on the standard canonical correlation approach although other estimators and in particular GMM are also possible (Anderson and Vahid 1998). Schleicher (2006) and Candelon, Hecq and Verschoor (2005) compare the behaviour of different estimators such as ML, canonical correlation, several GMM estimators and non parametric approaches for common cyclical features. The common cyclical feature test statistics can be implemented if $\hat{\delta}_3$ is a superconsistent estimate of $\delta_3$.

Schleicher (2006) proposes to test for codependent cycles using a maximum likelihood approach. Alternatively, information criteria can be used (see also Vahid and Issler 2002).

Finally, the following contributions are worth mentioning. Harding and Pagan (2006) deal with synchronized cycles and propose robust test to evaluate the presence of (un)synchronized cycles. Paruolo (2006) analyzes common cycles in I(2) VAR systems, while Johansen (2006b) proposes a novel asymptotic theory on cointegrating relations in I(2) models. See also Luginbuhl and Koopman (2004) and Morley (2006), where authors use a multivariate common converging trend-cycle decomposition and unobserved components models with common cycle restrictions, respectively.

An alternative to reduced form VAR modelling is the structural time series approach of Harvey (1989) and Harvey and Koopman (1997).

Common trends and common cycles in "structural time series" models (Carvalho, Harvey and Trimbur, in this volume)

In this paper, Carvalho, Harvey and Trimbur characterise common trends and common cycles in the structural time series models (STM) framework as an alternative to the large literature on common trends and common cycle features in reduced form VAR models. The authors review multivariate STMs and show how this framework describes trends, cycles and the interaction between these components in different series. Their approach provides a direct description of a series and yields reliable forecasts. Interactions between components in different series are easily captured by multivariate structural time series models (VECM models) and it is straightforward to impose common feature restrictions. The way in which trends are handled is highlighted by describing a structural time series model, as proposed in Carvalho and Harvey (2005), that allows convergence to a common growth
path. Post-sample data are used to test the forecasting performance of the model when fitted to income per head in US regions. Common cycles are then shown to be a special case of "similar cycles". That arises when the system of white noise processes that generates the similar cycles is singular. A test for common cycles is proposed, its asymptotic distribution is given and small-sample properties are studied by Monte Carlo experiments. The estimation and testing procedures are applied to US and Canadian GDP and to US regions, with special attention being paid to the implications of using higher-order cycles.

2.2 Common Seasonality

Seasonality, defined as "special annual dependence" (Sargent 2001), is another important feature of many economic time series. Seasonality is traditionnally modelled by means of seasonal dummy variables, unobserved periodic factors (e.g. within a structural time series model or an exponential smoothing model), or seasonal AR and/or MA polynomials. A more sophisticated model of seasonality is the "periodic model", which can be viewed as an ARMA model with varying parameters. In what follows, we briefly recall the concept of periodic integration and periodic cointegration, i.e. features and common features that describe periodic trends. Interested readers can refer to Franses (1996, 1998), Franses and Paap (2004), Ghysels and Osborn (2001) for a comprehensive presentation of the various issues related to seasonality.

A quarterly series $y_t$ follows a first-order periodic autoregression, $PAR(1)$, if it follows the representation $y_t = \phi_q y_{t-1} + \varepsilon_t$ with $t = 1, \ldots, T$ and $q = 1, \ldots, 4$, where $\varepsilon_t$ is white noise process, and where the parameter $\phi_q$ varies across seasons (Boswijk and Franses 1995). Let us define an annual process $Y_\tau$, where $\tau = 1, \ldots, T/4$, by stacking quarterly observations $y_t$ in column of vectors. $Y_\tau$ is thus the vector of quarters (VQ) process of $y_t$ from which it is evident that any periodic autoregression $y_t$ (time-varying model) implies a vector autoregression for $Y_\tau$ (constant parameter model).

**Definition 2** A process $y_t$ is periodically integrated of order 1, $PI(1)$, if the characteristic equation of the VAR model of $Y_\tau$ has exactly 1 unit root and all other roots lie outside the unit circle. If there are no unit roots, then $y_t$ is periodically stationary, $PI(0)$.
**Corollary 1** The condition of one unit root implies that the four components of \( Y_t \) have a single common trend and thus 3 cointegrating relationships.

Boswijk and Franses (1995) define the following concepts of periodic cointegration:

**Definition 3** Consider an \( n \)-dimensional vector of a quarterly process \( y_t \), with a VQ process integrated of order PI(1) and denote a set of seasonal dummies \( D_{st} \) with \( t = 1, \ldots, T \) and \( q = 1, \ldots, 4 \). Then \( y_t \) is said to be periodically cointegrated of order \((1, 1)\), if there exist \( n \times r \) matrices \( \delta_q \) of full rank \( (0 < r < n, q = 1, \ldots, 4) \) such that the VQ process of \( \sum_{q=1}^{4} D_{st} \delta_q y_t \) is stationary.


Although periodic models are flexible enough to incorporate seasonal characteristics of the series, they are not parsimonious because the number of parameters to estimate is a function of the sampling frequency and of the periodicity of observations. More parsimonious representations of periodic models have been derived by adapting the concept of serial correlation common features, as proposed in EK. Further, the current literature on periodic models has worked with the concepts of periodic integration and periodic cointegration, i.e. features and common features that describe periodic trends, but has not developed yet analogous concepts for periodic cycles.

**Common Periodic Correlation (Haldrup, Hylleberg, Pons, and Sanso, in this volume)**

In this paper, Haldrup, Hylleberg, Pons, and Sanso develop the concept of **common periodic correlation (CPC)** by adapting the weak and strong form features proposed by Hecq, Palm and Urbain (2006) to periodic models, and also by extending the non-synchronous approach in Cubadda and Hecq
(2001), and Cubadda (1999, 2001). The authors propose a multivariate representation of univariate and bivariate (possibly non-stationary) periodic models as a benchmark for imposing CPC feature restrictions in order to obtain parameter parsimony. CPCs are short-run common dynamic features that co-vary across the different days of the week and possibly also across weeks, and that can be common across different time series. It is also shown how periodic models can be used to describe interesting dynamic links in the interaction between stock and flow variables. The proposed modelling framework is applied to series of daily arrivals and departures in airport transits.

2.3 Common Structural Breaks

The econometrics of structural breaks has flourished since the seminal contributions by Rappoport and Reichlin (1989) and Perron (1989). [Banerjee and Urga (2005) and Perron (2006) offer the most recent comprehensive reviews on the topic.] Very little attention, instead, has been given to the evaluation of common structural breaks or co-breaking.

The concept of co-breaking was introduced by Hendry (1996) and Hendry and Mizon (1998). In this case, the feature is the location shift and co-breaking corresponds to the situation where some linear combinations of variables show no location shifts. If the location shifts occur at the exact same time periods in all variables, the cofeature combinations involve contemporaneous variables (contemporaneous mean co-breaking), and if common shifts occur at different time periods, the cofeature combinations involve current and lagged values (intertemporal mean co-breaking).

Following Hendry (1996) and Clements and Hendry (1998) co-breaking can be defined as the cancellation of location shifts across linear combinations of variables. We refer to location shifts as shifts in the unconditional expectations or shifts in the parameters of deterministic components.

To illustrate this point, let us consider the time series process $y_t$ for $t = 0, ..., T$ with

$$E(y_0) = k.$$

For $t > 0$,

$$E(y_t - k) = \mu_t \in \mathbb{R}^n$$

and when $\mu_t \neq 0$, then a location shift is said to occur in the process $y_t$.

**Definition 4** Let us now suppose there exists an $(n \times r)$ matrix $\delta_\pi$ of rank $r$ ($0 < r < n$) such that $\delta_\pi \mu_t = 0$. Contemporaneous mean co-breaking (CMC)
occurs when
\[ E(\delta_y^t - \delta_y^k) = \delta_y^t \mu_t = 0. \] (7)

Then the \( r \) linear transforms \( \delta_y^t \), of order \( (r \times 1) \), are independent of location shifts, thus cancelling any deviations between different time intervals. There is CMC of order \( r \) in the process \( y_t \) when the location shifts occur at the same point in time.

The rank \( r \) of \( \delta_y \) determines the number of co-breaking relationships while the linearly independent columns of \( \delta_y = (\delta_{y1}, ..., \delta_{yr}) \) are referred to as the co-breaking vectors.

**Definition 5** Consider the lag polynomial \( (n \times s) \) matrix \( \delta_y(L) \) of degree \( p > 0 \), such that \( \delta_y(L) = \sum_{i=0}^{p} \delta_{yi} L^i \). Intertemporal mean co-breaking (IMC) occurs when
\[ E(\delta_y^t(L)y_t - \delta_y^k(L)k) = \delta_y^t(L)\mu_t = 0. \] (8)

There is IMC when the mean location shifts occur at different times and are only eliminated in linear combinations of current and lagged values of variables, such as \( \sum_{i=1}^{p} \sum_{j=0}^{p} \delta_{yi} \cdot \mu_{i,t-j} \).

Condition (8) requires only that the reduced s dimensional system of \( p \) lags in \( n \) variables \( \delta_y^t(L)y_0 \) are independent of location shifts. The IMC captures breaks that caused a delayed location change or relationship change between variables.

A special case of CMC is known as cointegration cobreaking. Consider the cointegrated VECM representation:
\[ \Delta y_t = \gamma + \alpha(\delta_{\beta}y_{t-1} - \mu) + \epsilon_t \]

where the drift, \( \gamma \) and mean, \( \mu \), are subject to location shifts at time \( T \). Let the new shifted parameters be \( \gamma^* = \gamma + \theta \gamma \) and \( \mu^* = \mu + \theta \mu \) so that after the shift
\[ \Delta y_{T+1} = \gamma^* + \alpha(\delta_{\beta}y_T - \mu^*) + \epsilon_{T+1} \]
\[ = \gamma + \theta \gamma + \alpha(\delta_{\beta}y_T - (\mu + \theta \mu)) + \epsilon_{T+1} \]
\[ = \gamma + \alpha(\delta_{\beta}y_T - \mu) + \epsilon_{T+1} + [\theta \gamma - \alpha \theta \mu] \]
\[ = \Delta \bar{y}_{T+1} + [\theta \gamma - \alpha \theta \mu]. \]
The last representation is equal to the original one plus a composite shift in the parameters. In this case, we can construct an \( n \times r \) matrix \( \Pi \) of rank \( r \) \((0 < r < n)\) and form \( r \) linear combinations

\[
delta_n^r \Delta y_{T+1} = \delta_n^r \Delta \bar{y}_{T+1} + \delta_n^r [\theta \gamma - \alpha \theta \mu].
\]

**Definition 6** There is equilibrium-drift cobreaking if

\[
delta_n^r \theta \gamma = 0,
\]

while equilibrium-mean cobreaking exists if

\[
delta_n^r \alpha \theta \mu = 0.
\]

It can be showed that there is a close relationship between cointegration and co-breaking for some parameter changes, and that the common trends (between cointegrated variables) are equilibrium-mean cobreaking (i.e. eliminating shifts in mean) and the cointegration vector is equilibrium-drift co-breaking (i.e. eliminating shifts in drift).

The entire framework presented so far assumes knowledge that the location shifts happened at time \( T \). Various methods for dating the location shifts have been developed. Bai, Lumsdaine, and Stock (1998) developed methods for dating the location shifts in cointegrated systems. In particular, they construct confidence intervals for the "date of a single break" in multivariate time series.

**Co-breaking (Hendry and Massmann, in this volume)**

In this paper, Hendry and Massmann provide both a comprehensive review of old and new results on co-breaking. They establish a consistent terminology, collect theoretical results, delimit co-breaking to cointegration and common features, and review recent contributions to co-breaking regressions and the budding analysis of co-breaking rank. The authors discuss the importance of co-breaking for policy analysis, with special emphasis on impulse-response functions. A new procedure for co-breaking rank testing is presented, evaluated by means of Monte Carlo experiments, and illustrated using UK macro-economic data.
2.4 Factor analysis and principal components

Factor analysis and principal component analysis were originally developed to capture the main sources of variation and covariation among $n$ independent random variables in a panel framework. These methods were then extended by Brillinger (1969), and Geweke (1977) and Sargent and Sims (1977) into dynamic principal component analysis models and dynamic factor model respectively, able to predict the covariation in economic variables by few underlying latent factors. Although the two methods differ for small $n$, they give similar inferences as $n$ increases and gets large. Chamberlain and Rothschild (1983) categorised the dynamic models as exact and approximate dynamic factor models.

Since the factors are generally unobserved, they are usually estimated from sample covariance matrices using statistical techniques. There exist estimation procedures such as maximum likelihood and principal component analysis to estimate latent factors and their factor loadings. Though the maximum likelihood method is successful in estimating factors in low-dimensional systems, computational complexities arise in maximizing the likelihood functions over large number of variables as the cross-sectional dimensions increase. Dynamic principal component analysis however is easy to compute for higher dimensional panels and therefore has been widely used in the term structure literature.

Connor and Korajczyk (1986) prove the consistency of the latent factors estimated via principal component analysis when the cross-sectional dimension ($n$) is greater than the time-series dimension ($T$). However, from the standard factor analysis literature (see Anderson 2003) it is well-known that consistent estimation of latent factors cannot be achieved when either $n$ or $T$ is finite. Bai (2003) developed an inferential theory for large panels where the cross-sectional dependence is explicitly considered in the factor loadings representation.

Another important question concerns the number of factors in existence. Connor and Korajczyk (1993) estimated the number of factors using sequential limit asymptotics whereby $n$ tends to infinity first, followed by $T$. Cragg and Donald (1997) show that the BIC information criterion could be used to infer about the rank of a consistently estimated sample covariance matrix for $n$ and $T$ fixed. Stock and Watson (1999) assume that $\sqrt{n}/T$ goes to zero and estimate the number of latent factors. Bai and Ng (2002) consider the case of $n$ and $T \rightarrow \infty$ and develop a statistical procedure that consistently
estimates the number of factors from observed data. The paper asserts that the true number of factors $k$ can be estimated, for the stationary case, by minimizing one of the information criteria provided [see Bai and Ng 2002, pp. 201-202]. In the presence of nonstationary models, see Bai (2004).

More recent contributions to dynamic factor models are aimed at constructing indicator variables which reflect for instance the stage of the business cycle or which can be used as leading indicators to forecast important economic variables. In much of this literature, large sets of variables that are potentially useful for forecasting are used to construct indicator variables, see e.g. Stock and Watson (1999), Forni, Hallin, Lippi and Reichlin (2000). The latter consider a generalized factor model with infinite dynamics and non-orthogonal idiosyncratic components. The variables are required to have a bounded spectrum. They show that when there are $q$ dynamic factors as the number of variables [i.e. the cross-sectional dimension] and time tend to infinity, the first $q$ empirical dynamic principal component series of the variables converge to the factor space and the empirical counterpart of the projection of each variable on the leads and lags of these $q$ principal components converges to the common component of the variable. [Forni and Lippi (2001) provide a sound characterization theorem for the general dynamic factor model]

Interested readers can refer to Vahid (2006) for a comprehensive presentation of the procedure. See also Forni, Hallin, Lippi, Reichlin (2004, 2005) and Giannone, Reichlin and Sala (2005, 2006). Stock and Watson (1999) also consider dynamic factor models with infinite cross-sectional dimension. They allow for time-varying parameters and for the common component to be the projection of the variable on a finite number of leads and lags in the principal components. Recently, Aiolfi, Catao and Timmermann (2006) use a combination of the "static" approach by Stock and Watson (2002) and the "dynamic" one in Forni et al (2000), after some scaling, as proposed in Boivin and Ng (2006). Stock and Watson (2005) show that factors estimated via static principal components are highly correlated with factors estimated via weighted principal components. Finally, in a recent paper, Amengual and Watson (in this volume) propose an extension of the Bai and Ng (2002) estimator, valid for estimation of the number of static factors, to consistently estimate the number of dynamic factors.
Determining the Number of Underlying Factors in Macroeconomic Variables (Bai and Ng, in this volume).

In this paper, Bai and Ng propose a methodology to determine the number of shocks (dynamic factors) driving economic fluctuations without having to estimate the dynamic factors. In macroeconomic analysis, in general it is assumed that the number of shocks driving economic fluctuations, \( q \), is small. This assumption is left untested. The authors associate \( q \) with the number of dynamic factors in a large panel of data. They estimate a VAR in \( r \) static factors, where the factors are obtained by applying the method of principal components to a large panel of data. The eigenvalues of the residual covariance or correlation matrix are computed. They then test whether their eigenvalues satisfy an asymptotically shrinking bound that reflects sampling error. The procedure is applied to determine the number of primitive shocks in a large number of macroeconomic time series. An important result of this analysis is that the relationship between the dynamic factors and the static factors is precise.

2.5 Common ARCH Features

The presentation of features so far has been concerned with common feature analysis in the conditional mean of a set of variables. There is also a large body of literature on common features in conditional variance. Common factor ARCH models have been extensively studied, mainly with the aim of addressing the problem of large dimension of the parameter space in multivariate models of conditional second moments [see e.g. Engle, Ng and Rothschild (1990), and Engle and Marcucci (2006)].

Consider an \( n \)-dimensional (returns) process \( y_t \) such that \( y_t = (y_{1,t}, \ldots, y_{m,t})' \), with \( I_t(y) \) the information set of \( y \) at time \( t \)

\[
y_t \mid I_{t-1}(y) \sim iidD(\mu_t, H_t) \quad (9)
\]

with \( D(\mu_t, H_t) \) is a "unspecified" multivariate distribution with time-varying mean \( \mu_t \) (\( n \times 1 \)) and var-cov matrix \( H_t \) (\( n \times n \)). This formulation embodies all existing multivariate GARCH representations.

Let us define \( \varepsilon_t = y_t - \mu_t \). In general, a multivariate GARCH process can be represented as

\[
H_t = W + A(L) (\varepsilon_{t-1}\varepsilon_{t-1}') + B(L)H_{t-1}, \quad (10)
\]
where $A(L) = \sum_{i=1}^{q} A_i L^i$ and $B(L) = \sum_{j=1}^{p} B_j L^j$, and $L$ is the lag operator.

Various multivariate GARCH specifications (see the excellent survey by Bauwens, Laurent and Rombouts 2006) can be derived to make the formulations less complex and more parsimonious: the "vectorial ARCH (or Vech) representation" and the diagonal Vech representation of Bollerslev, Engle and Wooldridge (1988), a restricted version of the vectorial ARCH by assuming the matrices $A_i$ and $B_j$ diagonal; the "BEKK representation" of Engle and Kroner (1995); the multivariate GARCH with "conditional constant correlation" (CCC) of Bollerslev (1990); the "dynamic conditional correlation" (DCC) multivariate GARCH of Engle (2002) and Tse and Tsui (2002) and the various contributions generated by these papers. See inter alia Cajigas and Urga (2006), Palandri (2005), Engle and Colacito (2006). Of interest are also the so-called FLEX GARCH of Ledoit, Santa-Clara and Wolf (2003), and the decomposition of the conditional density into the product of the marginals and the copula proposed by Granger, Teräsvirta and Patton (2006).

An alternative approach to modelling multivariate conditional variances is the orthogonal GARCH model of Alexander (2001) in which the serial correlation feature in the second moment of a set of variables is assumed to be adequately characterised by GARCH specifications for the conditional variances of the first few principal components of these variables. Van der Weide (2002) argues that the rotation of variables suggested by ordinary principal component analysis (PCA) may not be the most appropriate for the goal of modelling the predictability of variances of a set of variables, and proposes a more general rotation which encompasses ordinary PCA as a special case. On a similar note, Perignon and Villa (2006) suggested modelling the time varying covariance matrix (volatility) of the common factors using a new methodology based on common principal component (CPC) analysis instead of the classical PCA.

Another parameterization of MGARCH commonly found in the literature is the Factor-GARCH (FGARCH) model of Engle, Ng, and Rothschild (1990), where a factor or few factors are responsible for the co-movement of all series,

$$
H_t = W + \sum_{k=1}^{K} \lambda_k \chi_k^{\prime} \left( \sum_{j=1}^{q} \alpha_{kj}^2 \varepsilon_{t-j}^{\prime} \varepsilon_{t-j} w_k + \sum_{j=1}^{p} \beta_{kj}^2 w_k^{\prime} H_{t-j} w_k \right), \quad (11)
$$
where $\lambda_k$ and $w_k$ are vectors of dimensions $(n \times 1)$, with the vector $\lambda_k$ representing the $k$-th factor loading and the scalar $w_k \varepsilon_t (= f_{k,t})$ the $k$-factor; $\alpha_{kj}$ and $\beta_{kj}$ are scalars. The model alleviates the over-parameterization of Vech and BEKK models. The main limitation of these models is that they are difficult to generalize to cases where assets have extra risk factors.

**Generalized orthogonal GARCH (Lanne and Saikkonen, in this volume)**

In this paper, Lanne and Saikkonen introduce a new kind of generalized orthogonal GARCH model that allows for a reduced number of conditionally heteroskedastic factors and, hence, idiosyncratic shocks. The authors propose a multivariate generalised orthogonal factor GARCH model more parsimonious and easier to estimate than the FGARCH model, and develop test procedures for checking the correctness of the number of factors. Parameter estimation method (maximum likelihood) and statistical procedures to specify an appropriate number of factors are proposed under both Gaussian and a mixture of Gaussian likelihoods. An interesting result is that some parameters of the conditional covariance matrix which are not identifiable under normality can instead be identified when the mixture specification is used. They also report an empirical application, where they model a system of exchange rate returns and test for volatility transmission.

**2.6 Factor Analysis on Realized Diffusive Volatility**

It is well established that factor models play an important role in modelling asset returns and their volatility. There is also a growing interest in the literature in realized volatility and in modelling infrequent jumps. The observed volatility of asset returns is associated with an underlying continuous-path volatility process and jumps in asset prices. The volume of the literature on the identification, estimation and testing of the significance of jumps has been growing recently (Barndorff-Nielsen and Shephard 2004, 2006, 2007, and Andersen, Bollerslev and Diebold 2005). In this case the feature is the jump, and the co-jumping (Barndorff-Nielsen and Shephard 2004) corresponds to the situation where some linear combinations of variables show no jump.

Stochastic volatility models have been developed in financial economics to account for the interdependence in time varying volatility. Several variation processes have been introduced to capture this phenomenon observed
in financial markets. Barndorff-Nielsen and Shephard (2007) provide a basis of the so-called "econometrics of arbitrage-free price processes" including variation processes, realized power variation, realized bipower variation, multivariate realized bipower variation, and jumps.

The volatility observed in data can be associated with a continuous-time stochastic volatility process and jumps evident in data. Empirical studies find that jump diffusion type models outperform other stochastic volatility models for pricing derivatives. Literature has also documented that jumps have very low persistence, meaning that one could measure jumps separately. Andersen, Bollerslev and Diebold (2005) assess the additional value in forecasting regressions by incorporating jump components into the realized quadratic variation and find highly significant jump coefficient estimates.

Let us suppose that \( y_t \) represents the daily return series, say \( r_t \). The jump component \( J_t \) can be consistently estimated as a difference between realized quadratic variation, defined as \( RQV_t = \sum_{j=1}^{M} r_{t,j}^2 \) with \( r_{t,j} = p_{t-1+\frac{j}{M}} - p_{t-1+\frac{j-1}{M}} \) where \( M \) refers to the sampling frequency, \( j \)th within-day return of day \( t \), and realized bipower variation, defined as \( RBPV_t = \frac{M}{2} \sum_{j=2}^{M} |r_{t,j}|r_{t,j-1}| \),

\[
RQV_t - RBPV_t = \sum_{j=1}^{N_t} C(j) = J_t
\]

where \( N \) is a finite counting process and \( C \) is the jump size. Huang and Tauchen (2005) report that an empirically more robust measure is given by the logarithmic ratio: \( J_t = \log RQV_t - \log RBPV_t \).

Testing procedures have been proposed to test for the presence of jumps. Barndorff-Nielsen and Shephard (2006) and Huang and Tauchen (2005) propose an approach based on bi-power variation measures, with noticeable applications in Andersen, Bollerslev and Diebold (2005), and Ghysels, Santa Clara and Valkanov (2006).

Common Features in the Conditional Variance in the Presence of Jumps (Anderson and Vahid, in this volume)

In this paper, Anderson and Vahid base their factor analysis on the approximate factor model developed by Chamberlain and Rothschild (1983), using the associated model selection criteria suggested by Bai and Ng (2002) to select the number of common factors. They also use the procedures discussed
in Barndorff-Nielsen and Shephard (2006a) to remove the jumps from the data. This procedure allows authors to focus on bi-power variation which provides a consistent estimate of integrated volatility in a continuous time model of the logarithm of the stock price in the presence of jumps. The authors argue that large jumps in asset prices, if untreated, can adversely affect the outcome of principal component based procedures that are used in determining the number of factors in the variances of a large set of asset returns. They suggest estimation procedures for approximate factor models that are robust to jumps when the cross-sectional dimension is not very large, and work with volatility measures that have been constructed in such a way that they contain no jump components. They develop multivariate factor models for forecasting volatility in Australian stocks. They report out of sample forecast analysis showing that multivariate factor models of volatility outperform univariate models, but that there is little difference between simple and sophisticated factor models.

3 FINAL REMARKS

The main challenge in editing this volume was to follow up the growing body of literature on the topics which lie under the common features umbrella. No doubt, econometric analysis will continue to benefit from methodological advancements in the field. We foresee exciting challenges ahead. First of all, it will be interesting to explore further estimation and testing procedures which allow, when relevant, to jointly identify the various common features, using an integrated approach rather than evaluating each feature individually. Second, more is expected on testing for the presence of more than one common (similar) cycle, especially in the structural time series approach. Third, issues related to the distribution of test statistics for periodic (multi)integration and (multi)cointegration, using more parsimonious periodic models, need to be addressed. Fourth, though there is a large body of literature "dealing with breaks", more work is expected to model and test for co-breaking both in time series and panel data. Fifth, the importance of modelling jumps in financial time series is now well established, and more contributions are expected in evaluating the information content of the presence of co-jumps and in developing appropriate testing procedures to detect them. Finally, the DCC approach has greatly contributed to solving the "dimensionality curse" in MGARCH models. However more challenges lie ahead with reference in
particular to the identification of appropriate multivariate (conditional) distributions (including conditional copula) of innovations, able to characterise financial time series better than the standard normal distribution, with obvious implications for estimation methods and testing procedures used.

Acknowledgments

It has been a pleasure to edit this volume. I greatly benefited from a fruitful collaboration with contributors and helpful advice from referees. I wish to thank Dennis Philip for preparing some preliminary material for this editorial. I am greatly in debt to Heather Anderson, Jushan Bai, Chris Brooks, Domenico Giannone, Andrew Harvey, David Hendry, Sharon Kozicki, Stepana Lazarova, Michael Massmann, Denise Osborn, Markku Lanne, Serena Ng, Dennis Philip, Lorenzo Trapani, Christoph Schleicher and Farshid Vahid for providing me with detailed and extremely helpful comments on a previous version of the paper. An anonymous referee provided useful comments and constructive suggestions to improve the paper. Responsibility for any remaining errors rests with me only.

Special thanks to the Editor, Torben G. Andersen, for his support and encouragement throughout the project. An Associate Editor offered valuable advice and suggestions. Thanks are also due to Alain Hecq who helped me during the reviewing process.

20 papers were submitted for this special volume and 6 were accepted for publication. Each paper was refereed anonymously by at least three referees, and in one case an Associate Editor was involved. To them my special thanks for their cooperation, valuable comments and helpful suggestions.

I would like to thank Cass Business School and the Centre for Econometric Analysis for the financial support provided to run the conference on "Common Features in London" held at Cass on the 16-17, December 2004 (http://www.cass.city.ac.uk/conferences/cfl/index.html) which generated this work.

I wish to dedicate this special issue of the Journal of Business & Economic Statistics to the memory of my mentor and friend Carlo Giannini of University of Pavia (Italy). Those who had the privilege to work with Carlo are missing his impeccable rigour and passionate dedication to econometric analysis.
References


[10] Bai, J., and S. Ng (in this volume), "Determining the Number of Primitive Shocks in Factor Models."


[34] Carvalho, V., Harvey, A., and Trimbur, T. (in this volume), "A Note on Common Cycles, Common Trends and Convergence."


