Testing for Common Autocorrelation in Data Rich Environments

Gianluca Cubadda
Università di Roma "Tor Vergata"
Italy

Alain Hecq
Maastricht University
The Netherlands

December 11, 2008

Abstract

This paper proposes a strategy to detect the presence of common serial correlation in large dimensional systems. We show that partial least squares can be used to consistently recover the common autocorrelation space. Moreover, a Monte Carlo study reveals that univariate autocorrelation tests on the factors obtained by partial least squares outperform traditional tests based on canonical correlation analysis. Some empirical applications are presented to illustrate concepts and methods.

JEL: C32

Keywords: Serial correlation common feature, reduced rank regression, partial least squares

1 Introduction

In the terminology of Engle and Kozicki (1993), a feature is a characteristic of a time series such that serial correlation, trend, seasonality, volatility, etc. Common features arise when variables have such features in common in the sense that there exist linear combinations of the series that fail to have the feature even though each of the series individually has the feature.

This paper focuses on the serial correlation feature among the elements of an n-dimensional stationary time-series y_t. In particular, we assume that y_t is generated by the Vector Auto-Regressive model of order p (VAR(p) hereafter)

$$\Phi(L)y_t = \varepsilon_t, \qquad t = 1, \dots, T \tag{1}$$

where $\Phi(L)$ is a p-order polynomial matrix such that $\Phi_0 = I_n$ and the roots of $|\Phi(L)|$ lie outside the unit circle, ε_t are i.i.d. $N_n(0, \Sigma_{\varepsilon\varepsilon})$, and, for simplicity, no deterministic elements are introduced.

The presence of serial correlation common features (SCCF hereafter) requires the existence of a full-rank $n \times s$ matrix δ $(s < n)$ such that $\delta' y_t = \delta' \varepsilon_t$ (Vahid and Engle, 1993). SCCF allows for rewriting the VAR (1) into the following Reduced Rank Regression (RRR henceforth) model

$$y_t = \delta_\perp \psi' w_{t-1} + \varepsilon_t, \tag{2}$$

where $w_{t-1} = [y'_{t-1}, \cdots, y'_{t-p}]'$ and ψ is a full-rank $np \times (n-s)$ matrix and $\delta_\perp$ is a full-rank $n \times (n-s)$ matrix such that $\delta'_\perp \delta = 0$.

It is worth noting that the $(n-s)$ common factors $F_t = \psi' w_t$ are responsible for all the predictable dynamics of the system (see, *inter alia*, Centoni *et al.*, 2007). Indeed, we have that

$$\mathrm{E}(y_t | w_{t-1}) = \delta_\perp F_{t-1}$$

which in turn implies

$$\mathrm{E}(\delta' y_t | w_{t-1}) = 0. \tag{3}$$

The above orthogonality condition reveals that either an instrumental variable approach or a canonical correlation analysis are appropriate statistical tools for conducting statistical inference on SCCF (Engle and Kozicki, 1993; Vahid and Engle, 1993).

However, when either the number of variables n or the number of lags p are too large relatively to the sample size T, it is convenient to resort to a weaker orthogonality condition than (3), namely

$$\mathrm{E}[\delta'_i y_t | (y'_{t-1}\delta_i, \cdots, y'_{t-p}\delta_i)'] = 0, \quad i = 1, 2, ..., s, \tag{4}$$

where $\delta = [\delta_1, \delta_2, \cdots, \delta_s]$ (see, e.g., Lucke, 1994). In particular, we propose to test for condition (4) by means of univariate tests for no autocorrelation of each $\delta'_i y_t$ having fixed δ to a consistent estimate of a base for the SCCF space.

It should be noticed from the outset that if (3) implies (4), the reverse does not necessarily hold. For instance, in the following bivariate VAR(1)

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ \phi_{21} & 0 \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix},$$

there is a single SCCF vector such as $\delta = [1, 0]'$ that implies $\mathrm{E}(\delta' y_t | \delta' y_{t-1}) = 0$. However we also have that $\mathrm{E}(\delta'_\perp y_t | \delta'_\perp y_{t-1}) = 0$ when $\mathrm{E}(\varepsilon_{1t}\varepsilon_{2t}) = 0$ and consequently there exists a second linear combination satisfying (4). Hence, our testing procedure may lack power against some alternatives. However, we think that cases where linear combinations of autocorrelated time series satisfy condition (4) but not condition (3) are of theoretical interest rather than of practical relevance.

A crucial role in our testing procedure is played by the estimator of the SCCF matrix. Although a Canonical Correlation Analysis (CCA, henceforth) provides the maximum likelihood estimator of δ, it may perform poorly with high-dimensional systems because inversions of large variance matrices are required. Hence, we explore the possibility of resorting to Partial Least Squares (PLS, henceforth) as an alternative to CCA. PLS, introduced by Wold (1985), are a family of multivariate techniques with the aim of maximizing the covariance between linear combinations of two variable sets, see, e.g., Rosipal and Krämer (2006) for a recent survey. Groen and Kapetanios (2008) have recently documented that PLS have superior forecasting performance than better known data-rich prediction methods as principal component and ridge regressions.

This paper is organized as follows. In Section 2, we discuss inference on SCCF, and we show that a form of PLS provides consistent estimates of a base of the SCCF space. In Section 3 we compare our procedures with traditional likelihood ratio tests for common features based on canonical correlation analysis by means of a Monte Carlo study. In Section 4 we present some empirical applications, and Section 5 concludes.

2 Common Serial Correlation, Partial Least Squares, and Canonical Correlation Analysis

In order to discuss statistical inference on the matrix δ, let us determine from the VAR under SCCF in (2) the covariance matrix $\mathrm{E}(y_t w'_{t-1})$

$$\Sigma_{yw} = \delta_\perp \psi' \Sigma_{ww},$$

where $\Sigma_{ww} = \mathrm{E}(w_{t-1} w'_{t-1})$. Since it follows that

$$\delta' \Sigma_{yw} = 0,$$

the matrix δ lies in the space generated by the eigenvector associated with the null eigenvalues of the symmetric semi-positive definite matrix $\Sigma_{yw}\Sigma_{wy}$.

In order to make the solution of this eigenvalue problem invariant to scale changes of individual elements of both y_t and w_t, we rather compute a base for the SCCF space as $[\nu_1^{PLS}, ..., \nu_s^{PLS}]$, where ν_i^{PLS} $(i = 1, 2, ..., s)$ is the eigenvector associated with the i-th smallest eigenvalue of the matrix

$$D_{yy}^{-1}\Sigma_{yw}D_{ww}^{-1}\Sigma_{wy},$$

with $\Sigma_{yy} = \mathrm{E}(y_t y_t')$, and D_{yy} and D_{ww} are diagonal matrices having the diagonal elements of, respectively, Σ_{yy} and Σ_{ww}. The solution of this problem is known in multivariate statistics as a form of PLS.[1]

A better known method to obtain the matrix δ is CCA, which can be seen as PLS after standardizing both the time-series vectors y_t and w_t. Indeed, since

$$\mathrm{E}(\Sigma_{yy}^{-1/2}y_t w_t'\Sigma_{ww}^{-1/2}) = \Sigma_{yy}^{-1/2}\Sigma_{yw}\Sigma_{ww}^{-1/2} = \Sigma_{yy}^{-1/2}\delta_{\perp}\psi'\Sigma_{ww}^{1/2},$$

the matrix δ lies in the space generated by $[\nu_1^{CCA}, ..., \nu_s^{CCA}]$, where ν_i^{CCA} $(i = 1, 2, ..., s)$ is the eigenvector associated with the i-th smallest eigenvalue of the matrix

$$\Sigma_{yy}^{-1}\Sigma_{yw}\Sigma_{ww}^{-1}\Sigma_{wy}.$$

Let $\widehat{\Omega}$ indicate the maximum likelihood estimator of a moment matrix Ω. Since the eigenvalues and eigenvector of a positive semi-definite matrix are continuous functions of that matrix (see, e.g., Magnus and Neudecker (1999)), by Slutsky's theorem the eigenvectors associated with the s smallest eigenvalues of both

$$\widehat{D}_{yy}^{-1}\widehat{\Sigma}_{yw}\widehat{D}_{ww}^{-1}\widehat{\Sigma}_{wy}, \tag{5}$$

and

$$\widehat{\Sigma}_{yy}^{-1}\widehat{\Sigma}_{yw}\widehat{\Sigma}_{ww}^{-1}\widehat{\Sigma}_{wy} \tag{6}$$

are consistent estimators of a base of the SCCF space. Moreover, we know that CCA provides the maximum likelihood estimator of δ under Gaussianity assumption (Anderson, 1984). However, since PLS require to invert diagonal matrices only, this method can provide estimates of δ (up to a normalization matrix) that are less disperse and more numerically stable when the dimension of w_t approaches the sample size T.

3 Test Statistics for Common Serial Correlation

In order to test for the presence of common cyclical features in a data-rich environment, we first consider the CCA framework. As shown by Anderson (1984), the likelihood ratio test for the null hypothesis that

---

[1] Notice that this solution, often referred as PLS-SB (see, e.g., Rosipal and Krämer (2006)), slightly differs from the original PLS algorithm as proposed by Wold (1985). Hereafter, we will refer to the considered solution simply as PLS.

there exist at least s SCCF vectors is based on the statistic

$$LR_s = -T\sum_{i=1}^{s} \ln(1 - \hat{\lambda}_i^{CCA}), \qquad s = 1, \ldots, n \tag{7}$$

where $\hat{\lambda}_i^{CCA}$ is the i-th smallest eigenvalue of the matrix (6). The test statistic (7) follows asymptotically a $\chi^2_{(v)}$ distribution under the null where $v = s \times np - s(n - s)$. The maximum likelihood estimator of the SCCF matrix is given by

$$\hat{\delta}^{CCA} = [\widehat{\nu}_1^{CCA}, \cdots, \widehat{\nu}_s^{CCA}],$$

where $\widehat{\nu}_i^{CCA}$ is the eigenvector associated with $\hat{\lambda}_i^{CCA}$ for $i = 1, 2, ..., s$.

It is clear that CCA may encounter numerical problems when n and/or p are large compared to the sample size T. Indeed, the matrices $\widehat{\Sigma}_{yy}$ and $\widehat{\Sigma}_{ww}$ could be (almost) numerically singular and estimates of small eigenvalues could be heavily biased. This is the reason why we also consider the PLS estimator of the SCCF matrix, denoted as

$$\hat{\delta}^{PLS} = [\widehat{\nu}_1^{PLS}, \cdots, \widehat{\nu}_s^{PLS}]$$

where $\widehat{\nu}_i^{PLS}$ is the eigenvector associated with the i-th smallest eigenvalue of the matrix (5) for $i = 1, 2, ..., s$.

In order to test for the presence of SCCF in systems where n and/or p are large, we consider tests based on the orthogonality condition (4) rather than (3). In particular, we test for the null hypothesis that $\delta_i' y_t$ is a white-noise for $i = 1, 2, ..., s$. In order to do that, we look at the Box-Pierce test statistics

$$Q_i^j = T\sum_{k=1}^{k} \hat{r}_{i,j}^2,$$

and their Ljung-Box refinements

$$Q_i^{j\prime} = T(T+2)\sum_{l=1}^{k} \frac{\hat{r}_{i,j}^2}{T-l},$$

where

$$\hat{r}_{i,j} = \left(\sum_{t=l+1}^{T} \frac{e_{t,i}^j e_{t-l,i}^j}{T-l}\right) \Bigg/ \left(\frac{1}{T}\sum_{t=1}^{T} {e_{t,i}^j}^2\right),$$

and $e_{t,i}^j = \hat{\delta}_i^{j\prime} y_t$, for $j = CCA, PLS$ and $i = 1, 2, ..., s$. Notice that, in principle, k should be chosen in accordance with the VAR order p and the cofeature rank s, see Cubadda *et al.* (2008b).

Both Q_i^j and $Q_i^{j\prime}$ follows asymptotically a $\chi^2_{(k)}$ distribution under the null. However, we should be careful about the size of the test when under the null hypothesis we assume the existence of more than one common feature vector, i.e. when $s > 1$. In this paper, we propose two different strategies to solve this problem.

In the first one, we use a simple Bonferroni bound approach and we apply a correction to keep the overall significance test at its nominal level. This amounts to compute

$$B_s^j = \max(Q_i^j) \text{ or } B_s^{j\prime} = \max(Q_i^{j\prime}), \qquad j = CCA, PLS$$

and to confront it with the critical level at $\frac{\alpha}{s}\%$ in the $\chi^2_{(k)}$.

In the second approach, based on Cubadda *et al.* (2008a), we test for the null of no autocorrelation on the aggregate $a_{t,s}^j = \sum_{i=1}^s e_{t,i}^j$ for $j = CCA, PLS$. Notice that under the null hypothesis $a_{t,s}^j$ is a white noise in large sample whereas $a_{t,s}^j$ does not converge to an innovation process when the SCCF rank is less than s. We denote the Box-Pierce and Ljung-Box test on $a_{t,s}^j$ as A_s^j and $A_s^{j\prime}$ respectively.

4 The Monte Carlo results

In order to investigate the small sample behavior of LR_s, B_s^j and $A_{s,t}^j$, we resort to a similar data generating process as in Cubadda *et al.* (2008a). In particular, we simulate a stationary VAR(1) with the following reduced rank structure

$$y_t = \alpha + \Phi_1 y_{t-1} + \varepsilon_t = \alpha + \delta_\perp C_1^{'} y_{t-1} + \varepsilon_t, \tag{8}$$

where $C_1^{'} = [0.5, -0.5, 0.5, -0.5, \cdots - 0.5, 0.75]$, $\delta_\perp' = [1, 1, \cdots, 1]$, α is a n dimensional vector of constant terms we generate from an uniform distribution on (0,1), and ε_t are i.i.d. $N_n(0, I_n)$.[2]

The system (8) has $s = n - 1$ SCCF relationships of the shape

$$\delta = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ -1 & -1 & \cdots & -1 & -1 \end{pmatrix}. \tag{9}$$

We use $M = 10.000$ replications and generate $T + 50$ observations with initial values set to zero and then the first 50 points are discarded to eliminate dependence from the starting conditions. We first consider $n = 9$ and 25 individuals for successively $T = 50, 200$ and 600 data points corresponding, say, to 50 annual, 200 quarterly and 600 monthly observations.

4.1 Results with known SCCF vectors

Table 1 compares the rejection frequencies from $s = 22$ to $s = n = 25$ of the canonical correlation LR_s tests (7) as well as the univariate tests B_s^0, $B_s^{0\prime}$, A_s^0 and $A_s^{0\prime}$, where the superscript "0" indicates that linear combinations with known coefficients of $[\delta, \delta_\perp]$ (9) are used in the testing procedures.[3] This allows

---

[2] Alternative choices of $\Sigma_{\varepsilon\varepsilon}$ do not significantly alter the results.

[3] Rejection frequencies for $s = 1$ to 21 are not reported to save space.

Table 1: Size and power of common features tests statistics

		$k=2$				$k=5$			
	$T\backslash s=$	22	23	24	25	22	23	24	25
B_s^0	50	4.14	4.18	4.29	96.1	9.98	10.1	10.3	93.2
	200	4.09	4.07	4.08	100	6.15	6.19	6.25	100
	600	4.34	4.3	4.24	100	5.25	5.26	5.22	100
$B_s^{0\prime}$	50	6.13	6.07	6.19	96.8	16.7	17	17.3	94.9
	200	4.6	4.6	4.52	100	7.17	7.25	7.28	100
	600	4.48	4.43	4.4	100	5.48	5.51	5.58	100
A_s^0	50	4.45	4.60	4.45	96.40	5.92	6	5.96	93.7
	200	4.98	4.95	5	100	5.14	5.22	5.2	100
	600	4.6	4.55	4.5	100	4.8	4.82	4.77	100
$A_s^{0\prime}$	50	5.75	5.69	5.60	95.80	8.3	8.48	8.57	94.9
	200	5.26	5.21	5.24	100	5.69	5.76	5.91	100
	600	4.73	4.64	4.6	100	4.99	4.98	5.04	100

Note: The columns $s \leq 24$ gives the empirical size while the column $s = 25 = n$ stands for the power. The DGP has 24 SCCF vectors. B_s^0, $B_s^{0\prime}$, A_s^0 and $A_s^{0\prime}$ are tests for the null of no autocorrelation in the linear combinations $[\delta' y_t, \delta'_\perp y_t]$ with the true values of the vectors $[\delta, \delta_\perp]$ in DGP (8).

to compare the behavior of the different tests proposed in the previous section when there is not estimation error.

The frequency for $s = 25$ is the power (size unadjusted) of the test, $s = 22, 23, 24$ is the empirical size, i.e. the frequency with which 24 vectors are not rejected. From Table 1 we see that the behavior of univariate tests for the null of no autocorrelation with the known combinations is very good as it might be expected with the exception in some circumstances of the Ljung-Box test for small sample sizes and large k.

4.2 Results with estimated SCCF vectors: comparison of CCA and PLS

Results from Table 1 refer to an unfeasible testing procedure in the sense that we use the true common feature vectors to construct the combination that are used for univariate autocorrelation tests.

Let us now compare the results of test LR_s with those of tests based the estimated combinations obtained from either CCA or PLS. To save space, we only report the rejection frequencies for $n = 9$ series for $s = 8$ and $s = 9$, namely the empirical size and the empirical power of these tests. We choose a smaller dimension than the one from Table 1 because we wish to illustrate that, even with this relatively small number of series, the CCA framework already shows some large size distortions. Given the findings of Table 1 we also focus on the Box-Pierce test instead of its modified Ljung-Box version.[4]

From Table 2 we see that it is only with a very large sample size such as $T = 600$ and $k = 2$ that one

---

[4] All results are available upon request.

Table 2: Size and power of common features tests statistics

		$k=2$		$k=5$	
	$T \backslash s =$	8	9	8	9
LR_s	50	93.2	100	100	100
	200	17.3	100	68.8	100
	600	7.97	100	16.2	100
B_s^{CCA}	50	10.2	97.3	38.4	54.7
	200	12.7	100	12.2	100
	600	14	100	13.4	100
B_s^{PLS}	50	10.2	97.9	16.4	95.7
	200	14.6	100	13.40	100
	600	15.7	100	11.7	100
A_s^{CCA}	50	3.75	7.54	8.82	10.9
	200	3.48	35.3	4.49	31.30
	600	3.25	62.2	4.20	60.9
A_s^{PLS}	50	5.17	95.5	6.19	93.1
	200	5.28	100	5.40	100
	600	5.04	100	5.24	100

Note: The column $s \leq 8$ gives the empirical size while the column $s = 9 = n$ stands for the (size non-adjusted) power. The DGP has 8 SCCF vectors.

observes roughly correct empirical size with the LR_s test. In other cases the frequency with which the null $s = 8$ is rejected is very high, e.g. 93% with $T = 50$ and $k = 2$.

Turning to autocorrelation tests on estimated linear combinations, it clearly emerges that tests based on PLS give better results in terms of size distortion and power than tests based on CCA. The best strategy is A_s^{PLS}. For this procedure, there are negligible size distortions and the power if high even in small samples. The results of B_s^{PLS} are a bit disappointing, especially when one compares them with the case of known common feature vectors. Interestingly enough however, alternative identifications of the SCCF vectors give different results but we leave this issue for further investigations. For instance, normalizing the estimated vectors such that $\widetilde{\delta}' = [I_s, \widetilde{\delta}_{s\times(n-s)}]$ provide tests with rejections frequencies that are very close to those obtained with known SCCF vectors.

We finally simulate series from the same DGP as in (8) but with $n = 25$ series. Only A_s^{PLS} and B_s^{PLS} are reported in Table 3 for the size $s = 24$ and the empirical power $s = 25$. The results confirm what we observed so far because we correctly detect with A_s^{PLS} the presence of 24 common feature vectors even with this large set of series. There is obviously a small size distortion for $T = 50$ and $k = 8$ (the empirical size is 7% in this worse case) but this case represents a situation in which CCA would not even be feasible in a system of 125 regressors for each variable.

We would consequently recommend to add the A_s^{PLS} statistics, namely a Box-Pierce test on the aggregate

Table 3: Size and power of common features tests statistics

		$k=2$		$k=5$		$k=8$	
	$T\backslash s=$	24	25	24	25	24	25
B_s^{PLS}	50	12.0	96.40	42.2	94.9	64.3	95.6
	200	22.60	100	30	100	34.9	100
	600	27.10	100	23.9	100	22.3	100
A_s^{PLS}	50	5.20	98.6	5.72	96.70	7.06	95.50
	200	5.03	100	5.31	100	5.36	100
	600	4.75	100	5.05	100	5.22	100

Note: The DGP has 24 SCCF vectors within 25 series. Hence the column $s = 24$ gives the empirical size while the column $s = 25 = n$ stands for the power (size unadjusted).

of the PLS factors, to the traditional toolkit for common features analysis.

5 Empirical Examples

This section presents a couple of applications, in which we apply both LR_s and A_s^{PLS} tests to some interesting empirical issues. Particularly, we first examine the serial correlation properties of revision errors based on the "preliminary" data releases of the EU12 industrial production index. Second, we analyze if the dynamics of the US economy become less predictable during the "Great Moderation".

5.1 Real time data

Data currently produced by statistical offices typically undergo a recurrent revision process resulting in different releases of the same phenomenon. Indeed, the nature of the statistical system imposes that data are improved when new important information becomes available (e.g. new input output matrices, survey or census data, etc.) or because of changes in definitions or statistical methods (e.g. chain linked index replacing a constant price index, new base year, updated seasonal factors, etc.). Consequently a database consists in general of vintages of the major macroeconomic data available in real time.

Let us denote $x_{t,t+v}$ the point estimate for x_t published in $t+v$ with $v \geq 0$. Collecting the whole series for $t = 1...T$ we have what is called a diagonal of the real time data matrix (Croushore and Startk, 2001; Croushore, 2008; Hecq and Mazzi, 2008). We can look at the different diagonals denoted $X_{t,t+v}$ and for instance at the next vintage diagonal $X_{t,t+v+1}$. The problem to know that the data revision process brings news or noise in the data depends on whether $X_{t,t+v+1} - X_{t,t+v}$ is serially correlated or a white noise (see, e.g., Croushore 2008).

We have considered vintage series for the EU12 industrial production index. Data have been rebased (see Hecq and Mazzi, 2008, for details). The 16 series are in months from $X_{t,t+3}$ to $X_{t,t+18}$ and $T = 58$.

Table 4: Number of SCCF's (with a significance level of 0.05) on 16 diagonal vintages

	$k=2$	$k=4$	$k=8$
LR_s	0	$-$	$-$
A_s^{PLS}	8	7	9

Note: The symbol "$-$" indicates that CCA is unfeasible due to lack of degrees of freedom.

The results, reported in Table 4, show that LR_s and A_s^{PLS} provide different results. The LR_s test is feasible only for $k = 2$ and it provides no evidence of SCCF, whereas the A_s^{PLS} tests suggests that approximately the half of the PLS linear combinations are white noises. Definitively, this means that there exist some genuine new information coming during the revision process.

5.2 Great moderation

Starting from the seminal paper of McConnell and Perez-Quiros (2000), a large of body of literature documented a clear decline in macroeconomic volatility in the US since the middle of the 1980's. This phenomenon, known as "the Great Moderation" has two alternative explanations. On one hand, it is attributed to a decline in the volatility of shocks ("good luck" hypothesis). On the other hand, the Great Moderation is seen as the result of a structural break in the shock propagation mechanism, mainly connected to change in the US monetary policy, which became more oriented towards inflation control since the nomination of Paul Volcker as the head of the Federal Reserve Bank ("good policy" hypothesis).

A possible way of discriminating between these two hypotheses is to check if the Great Moderation is associated with a change in the relative predictability of the key macro variables. Indeed, a change in the unconditional variance of a time series has no implications on its degree of predictability. As pointed out by D'Agostino *et al.* (2006), the accuracy of professional forecasts has dramatically worsened during the Great Moderation period. Following Giannone *et al.* (2008), we tackle this issue in a large dimensional framework. Particularly, we consider a system with 18 variables, which are listed in Table 5, along with their transformations. The data are observed at the monthly frequency for the period 1959.01-2003.12

In Table 6 we report the results of the application of both LR_s and A_s^{PLS} tests for the samples 1959–1983 (pre–Great Moderation) and 1984–2003 (Great Moderation). The two tests provide opposite conclusions. According to the LR_s tests, we even find less evidence of SCCF's in the Great Moderation period, whereas the A_s^{PLS} tests point out that the number of PLS factors with no autocorrelation has almost doubled in the second sub-sample. In view of the results of the Monte Carlo experiment in Section 4, we conclude that the evidence based on the LR_s statistics is likely to be spurious, since these tests suffer of sever size distortion in large dimensional systems. Instead, the A_s^{PLS} tests confirm previous findings in the literature.

Table 5: List of 18 US economic indicators

Variable	Transformation
Real personal income	$(1-L)\ln$
Real consumption	$(1-L)\ln$
Industrial production index	$(1-L)\ln$
Capacity Utilization	$(1-L)$
Unemployment rate	$(1-L)$
Average weekly hours	$(1-L)\ln$
NAPM inventories index	none
M2 (1996 prices)	$(1-L)\ln$
Non-borrowed reserves	$(1-L)^2\ln$
Standard & Poor's common stock price index	$(1-L)\ln$
Interest rate spread: 1 year	none
Interest rate spread: 5 years	none
Interest rate spread: 10 years	none
Producer price index	$(1-L)^2\ln$
NAPM commodity prices index	none
Consumer prices index	$(1-L)^2\ln$
Personal consumption expenditure deflator	$(1-L)^2\ln$
Average hourly earnings	$(1-L)^2\ln$

Table 6: Number of SCCF's on 18 US economic indicators

		$k=2$	$k=4$	$k=8$
1959:1983	LR_s	3	1	0
	A_s^{PLS}	2	2	1
1984:2003	LR_s	1	0	0
	A_s^{PLS}	3	4	2

6 Conclusion

This paper has illustrated that univariate autocorrelation tests on the PLS factors are helpful tools for testing the null hypothesis of common autocorrelation in a data rich environment.

A Monte Carlo study has shown that the new testing procedures outperform traditional likelihood ratio tests when the system dimension is large. A couple of empirical applications has illustrated the practical value of the proposed procedures.

These results leave room for several interesting developments, such as the analysis of common trends and forecasting in large dimensional systems with reduced rank structures. These issues are in our research agenda.

References

[1] ANDERSON, T.W. (1984), *An Introduction to Multivariate Statistical Analysis*, 2nd Ed. John Wiley & Sons.

[2] CENTONI, M., CUBADDA, G., AND A. HECQ (2007), Common Shocks, Common Dynamics, and the International Business Cycle, *Economic Modelling,* 24, 149-166.

[3] CROUSHORE, D. (2008), Frontiers of Real-Time Data Analysis, *Federal Reserve Bank of Philadelphia Working Papers Series,* 08-4

[4] CROUSHORE, D. AND T. STARTK (2001), A real-time data set for macroeconomists, *Journal of Econometrics,* 105, 111-130.

[5] CUBADDA, G., HECQ, A. AND F.C. PALM (2008A), Macro-panels and Reality, *Economics Letters,* 99, 537-540.

[6] CUBADDA, G., HECQ, A. AND F. PALM (2008B), Sudying Co-movements in Large Multivariate Models Without Multivariate Modelling, forthcoming in *Journal of Econometrics.*

[7] D'AGOSTINO, ANTONELLO, DOMENICO GIANNONE, AND PAOLO SURICO (2006), (Un)Predictability and macroeconomic stability, *European Central Bank Working Paper Series,* 605.

[8] ENGLE, R.F. AND S. KOZICKI (1993), Testing for Common Features (with comments), *Journal of Business and Economic Statistics*, 11, 369-395.

[9] GIANNONE, D., LENZA, M. AND L. REICHLIN (2008), Explaining The Great Moderation: It Is Not The Shocks, *Journal of the European Economic Association,* 6, 621-633

[10] GROEN, J.J., AND G. KAPETANIOS (2008), Revisiting Useful Approaches to Data-Rich Macroeconomic Forecasting, *Federal Reserve Bank of New York Staff Report*, 327.

[11] HECQ, A, AND G. L. MAZZI (2008), Advanced Revision Analysis for Economic Time Series and Their Role for Improving Forecast Accuracy." *Maastricht University Working paper*, June 2008.

[12] LUCKE, B. (1994), Analysis of West German Macroeconomic Data using Common Trends and Common Cycles, *Free University of Berlin Discussion Paper*, 3.

[13] MAGNUS J.R. AND H. NEUDECKER (1999), *Matrix Differential Calculus with Applications in Statistics and Econometrics*, 2nd Edition, John Wiley.

[14] MCCONNELL, M.M. AND G. PEREZ-QUIROS (2000), Output fluctuations in the United States: What has changed since the early 1980's, *American Economic Review*, 90, 1464-1476.

[15] ROSIPAL, R. AND N. KRÄMER (2006), Overview and Recent Advances in Partial Least Squares, in *Subspace, Latent Structure and Feature Selection*, Springer, Berlin, 34-51.

[16] VAHID, F. AND R.F. ENGLE (1993), Common Trends and Common Cycles, *Journal of Applied Econometrics,* 8, 341-360.

[17] WOLD, H. (1985), Partial least squares, in S. Kotz and N.L. Johnson (eds), *Encyclopedia of the Statistical Sciences*, John Wiley & Sons, 6, 581–591