Dynamic specification tests for factor models

Gabriele Fiorentini
<fiorentini@ds.unifi.it>
*Universita di Firenze*

Enrique Sentana
<sentana@cemfi.es>
*CEMFI*

Conference in honour of Andrew Harvey
Oxford Man Institute, June 30th, 2012
Acknowledgement

- It would have been virtually impossible to write this paper without Andrew Harvey’s contributions.

- We have borrowed many results from his books, articles and even unpublished papers, some of which are at risk of being lost forever in our digital world.

- We have also benefited from the work of his students and co-authors, many of whom are present.

- Therefore, it is an honour for me to present this paper in front of this audience.

- Obviously, we are fully responsible for any errors or misinterpretations that it may contain.
Outline of the presentation

1. Introduction

2. Serial correlation tests
   - Time domain
   - Frequency domain

3. Conditional heteroskedasticity tests

4. Non-normal distributions

5. Empirical application

6. Conclusions
Motivation

- There is a long tradition of factor or multi-index models in finance, where they were originally introduced to simplify the computation of the covariance matrix of returns in a mean-variance portfolio allocation framework.

- The common factors usually correspond to unobserved fundamental influences on returns, while the idiosyncratic factors reflect asset specific risks.

- The concept of factors plays a crucial role in the mutual fund separation theory, of which the standard CAPM is a special case, and the Arbitrage Pricing Theory.
Static factor models

\[ y_t = \pi + cf_t + v_t, \]

\[ \begin{pmatrix} f_t \\ v_t \end{pmatrix} \mid I_{t-1}, \theta_s \sim N \left[ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & \Gamma \end{pmatrix} \right] \]

- \( y_t \): \( N \times 1 \) vector of observable variables,
- \( \pi \): \( N \times 1 \) vector of constant conditional means,
- \( f_t \): unobserved common factor with \( V(f_t|I_{t-1})=1 \),
- \( c \): \( N \times 1 \) vector of factor loadings,
- \( v_t \): \( N \times 1 \) vector of idiosyncratic noises conditionally orthogonal to \( f_t \),
- \( \Gamma \): \( N \times N \) diagonal p.s.d. matrix of constant idiosyncratic variances,
- \( I_{t-1} \): information set that contains lagged values of \( y_t \) and \( f_t \),
- \( \theta_s = (\pi', c', \gamma')' \), with \( \gamma = vecd(\Gamma) \).

Our assumptions trivially imply that:

\[ y_t|I_{t-1}; \theta_s \sim N[\pi, \Sigma(\theta_s)], \]

\[ \Sigma(\theta_s) = cc' + \Gamma. \]
Static factor models

- The Kalman or Wiener-Kolmogorov filters yield the basic ingredients for our tests. Since the model is static:
  \[
  E \left( \begin{array}{c} f_t \\ v_t \end{array} \mid Y_t; \theta_s \right) = \begin{bmatrix} f_{kt}(\theta_s) \\ v_{kt}(\theta_s) \end{bmatrix} = \begin{bmatrix} c'\Sigma^{-1}(\theta_s)(y_t - \pi) \\ \Gamma \Sigma^{-1}(\theta_s)(y_t - \pi) \end{bmatrix}.
  \]

- The corresponding mean square errors are:
  \[
  \begin{bmatrix}
  \omega_{kt}(\theta_s) & -c'\omega_{kt}(\theta_s) \\
  -c\omega_{kt}(\theta_s) & ccc'\omega_{kt}(\theta_s)
  \end{bmatrix} = \begin{bmatrix}
  1-c'\Sigma^{-1}(\theta_s)c & -c'\Sigma^{-1}(\theta_s)\Gamma \\
  -\Gamma \Sigma^{-1}(\theta_s)c & \Gamma - \Gamma \Sigma^{-1}(\theta_s)\Gamma
  \end{bmatrix},
  \]
  which is a matrix of rank 1.

- The unconditional covariance matrix of \( f_{kt}(\theta_s) \) and \( v_{kt}(\theta_s) \) is also singular (rank \( N < N + 1 \)) because \( y_t = \pi + cf_{kt}(\theta_s) + v_{kt}(\theta_s) \).

- The asymptotic distribution of the Gaussian maximum likelihood estimators of the static model parameters is straightforward:
  \[
  \sqrt{T}(\bar{\theta}_s - \theta_{s0}) \rightarrow N\left[0, \mathcal{I}_{\theta_s}^{-1}(\theta_{s0})\right],
  \]
  where \( \mathcal{I}_{\theta_s}(\theta_{s0}) \) is the information matrix.
We initially develop tests of first order serial correlation in the common and idiosyncratic factors using a time domain approach.

Our model under the alternative is:

\[
\begin{align*}
    y_t &= \pi + cx_t + u_t \\
    x_t &= \rho x_{t-1} + f_t \\
    u_t &= \text{diag}(\rho^*)u_{t-1} + v_t
\end{align*}
\]

\[
\begin{pmatrix} f_t \\ v_t \end{pmatrix} | I_{t-1}, \theta \sim N \left[ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & \Gamma \end{pmatrix} \right]
\]

It reduces to our baseline specification under the null hypothesis that \( H_0 : \rho^+ = 0 \), where \( \rho^+ = (\rho, \rho^*)' \).

Otherwise, it has the autocorrelation structure of a VARMA(2,1).

This model has become increasingly popular in macroeconomic applications, but it is not widely used in finance.
Proposition 2

Let

\[ \bar{G}_{f_kf_k}(j) = \frac{1}{T} \sum_{t=1}^{T} f_{kt}(\theta_s)f_{kt-j}(\theta_s) \]

\[ \bar{G}_{v_kv_k}(j) = \frac{1}{T} \sum_{t=1}^{T} v_{kt}(\theta_s)v'_{kt-j}(\theta_s) \]

denote the \( j^{th} \) sample autocovariances of the static Kalman filter estimators of the common and specific factors.

1. Under the null hypothesis \( H_0 : \rho^\dagger = 0 \)

\[
LM_{AR(1)} = T \cdot (\bar{G}_{f_kf_k}(1), vecd'[\bar{G}_{v_kv_k}(1)\Gamma^{-1}) \\
\times \mathcal{I}_{\rho^\dagger \rho^\dagger}(\theta_{s0}) (\bar{G}_{f_kf_k}(1), vecd'[\bar{G}_{v_kv_k}(1)\Gamma^{-1})']
\]

converges to a \( \chi^2 \) with \( N + 1 \) degrees of freedom as \( T \to \infty \), where \( \mathcal{I}_{\rho^\dagger \rho^\dagger}(\theta_{s0}) \) is the relevant block of the information matrix.

2. This asymptotic null distribution is unaffected if we replace \( \theta_{s0} \) by its Gaussian maximum likelihood estimator \( \bar{\theta}_s \).
Time domain AR(1) tests in common and specific factors

- We can re-write this score test as a joint moment test based on the $N + 1$ orthogonality conditions that we would use to test for first order serial correlation if we treated $f_{kt}(\theta_s)$ or $v_{kit}(\theta_s)$ as the series of interest:

  \[
  E[f_{kt}(\theta_s)f_{kt-1}(\theta_s)|\theta_s,0] = 0, \\
  E[\gamma_i^{-1}v_{kit}(\theta_s)v_{kit-1}(\theta_s)|\theta_s,0] = 0 \quad (i = 1, \ldots, N).
  \]

- These moment conditions closely resemble the orthogonality conditions that we would use to test for first order serial correlation if we could observe all the latent variables:

  \[
  E(f_t f_{t-1}|\theta_s,0) = 0, \\
  E(\gamma_i^{-1}v_{it} v_{it-1}|\theta_s,0) = 0 \quad (i = 1, \ldots, N).
  \]

- Our joint tests, though, take into account the non-diagonal and singular nature of the covariance matrix of $f_{kt}(\theta_s)$ and $v_{kt}(\theta_s)$, even though the latent variables themselves are uncorrelated.
Other tests for multivariate serial correlation

- We can compare our test of serial correlation in common and specific factors to Hosking’s (1981) multivariate autocorrelation test.
- In the first order case, we can interpret his proposal as a test of the null hypothesis of lack of serial correlation against an unrestricted $\text{VAR}(1)$ model (as in Hendry (1971), Guilkey (1974) and Harvey (1982)).
- It is based on the first autocovariance matrix of the observed variables:
  \[
  \bar{G}_{yy}(1) = \frac{1}{T} \sum_{t=1}^{T} (y_t - \bar{\pi})(y_{t-1} - \bar{\pi})'
  \]
  where $\bar{\pi}$ is the sample mean.
- Under the null of multivariate white noise, this statistic is distributed as a $\chi^2$ with $N^2$ degrees of freedom as $T$ goes to infinity.
- The number of degrees of freedom is an order of magnitude larger than in our test because it does not exploit the strong cross-sectional dependence of returns, which reduces power.
- It also requires $T$ much larger than $N^2$ for the asymptotic distribution to be reliable in finite samples.
Power comparison: AR tests

- We compare the power of our LM tests, Hosking’s original test, a diagonal version, and a standard univariate first-order serial correlation test applied to the Equally Weighted Portfolio (EWP).
- To do so, we consider a non-exchangeable single factor model of the form:

\[
\begin{align*}
y_{it} &= \pi_i + c_i x_t + u_{it} \quad (i = 1, \ldots, 5) \\
x_t &= \rho x_{t-1} + (1 - \rho^2) f_t \\
u_{it} &= \rho^*_i u_{i,t-1} + (1 - \rho^*_{i}^2)v_{it}
\end{align*}
\]

- \( \pi = (0.5, 0.4, 0.5, 0.4, 0.5) \),
- \( c = (5, 4, 5, 4, 5) \),
- \( \gamma \propto (5, 9, 5, 9, 5) \),
- \( \rho^*_i = \rho^* \ \forall i \).
- We evaluate power against compatible sequences of local alternatives of the form \( \rho^{\dagger}_{0T} = \rho^{\dagger}_0 / \sqrt{T} \).
Power comparison: $\rho^* = 1.5\rho$, baseline signal to noise ratio

Power of mean dependence tests versus local alternatives (common & specific)

- Joint LM
- LM on common factor
- LM on Specific factors
- Hosking
- Diagonal Hosking
- EWP

Gabriele Fiorentini Enrique Sentana ()
Extensions

- MA processes for the latent variables, which turn out to be locally equivalent alternatives to AR processes, and therefore lead to numerically identical statistics.

- AR(h) tests, with or without distributed-lag or panel data type restrictions on the AR coefficients.

- Models with multiple factors, in which case we derive tests that are numerically invariant to rotations of the common factors.
Frequency domain tests for common and specific factors

- An alternative way to characterise a dynamic factor model is in the frequency domain.

- Assuming the stationarity conditions $|\rho| < 1$ and $|\rho_i^*| < 1$ hold $\forall i$, the spectral density matrix under the alternative will be given by

$$g_{yy}(\lambda) = cc'g_{xx}(\lambda) + g_{uu}(\lambda),$$

$$g_{xx}(\lambda) = \frac{1}{1 + \rho^2 - 2\rho \cos(\lambda)},$$

$$g_{uiu_i}(\lambda) = \frac{1}{1 + \rho_i^{*2} - 2\rho_i^* \cos(\lambda)},$$

which inherits the exact single factor structure of the unconditional covariance matrix $\Sigma(\theta_s)$.

- Importantly, $g_{yy}(\lambda)$ is not only positive semidefinite Hermitian but also symmetric, which simplifies the calculation of the log-likelihood function in the frequency domain.
Frequency domain tests for common and specific factors

- Let $\lambda_j = 2\pi j / T$ denote the usual Fourier frequencies. Ignoring the complex elements of the periodogram matrix

  $$
  \bar{g}_{yy}(\lambda) = \frac{1}{T} \sum_{t=1}^{T} \sum_{s=1}^{T} (y_t - \pi)(y_s - \pi)' e^{-i(t-s)\lambda},
  $$

  the (spectral) log-likelihood function is

  $$
  - \frac{NT}{2} \ln(2\pi) - \frac{1}{2} \sum_{j=0}^{T-1} \ln |g_{yy}(\lambda_j)| - \frac{1}{2} \sum_{j=0}^{T-1} \text{tr} \left\{ g_{yy}^{-1}(\lambda_j) \Re [\bar{g}_{yy}(\lambda_j)] \right\}
  $$

- The scores with respect to $\rho$ and $\rho_i^*$ evaluated under the null are

  $$
  \sum_{j=0}^{T-1} \cos(\lambda_j) \left[ \bar{g}_{x_ktx_k}(\lambda_j) - c' \Sigma^{-1}(\theta_s) c \right] \propto \sum_{j=1}^{T} x_{kt}(\theta_s)x_{kt-1}(\theta_s),
  $$

  $$
  \sum_{j=0}^{T-1} \cos(\lambda_j) \left[ \bar{g}_{u_kiu_k}(\lambda_j) - \sigma^{ii}(\theta_s) \gamma_i^2 \right] \propto \sum_{j=1}^{T} u_{kit}(\theta_s)u_{kit-1}(\theta_s),
  $$

  which coincide with the univariate scores if we treated $x_{kt}(\theta_s)$ and $u_{kit}(\theta_s)$ as if they were observed.
An important advantage of the frequency domain approach is that it allows us to easily deal with more complex situations.

Specifically, suppose that we take as our new null hypothesis the factor model with AR(1) dynamics in the latent variables that we have been considering as the alternative so far, and as our new alternative a model with a common factor that follows an AR(2) process.

Although a Lagrange Multiplier test of the new null hypothesis in the time domain is conceptually straightforward, the algebra is incredibly tedious and the recursive scores difficult to interpret.

In contrast, the frequency domain scores remain remarkably simple.

This is particularly true if we parametrise the dynamics of the latent variables in terms of partial autocorrelations, as in Quenouille (1949).

In the AR(2) case:

\[ x_t = \tau_1 (1 - \tau_2)x_{t-1} + \tau_2 x_{t-2} + f_t. \]
The score with respect to the second partial autocorrelation coefficient of the common factor, $\tau_2$, under the AR(1) null is

$$\sum_{j=0}^{T-1} \left[ \tau_1^2 - 2\tau_1 \cos(\lambda_j) + \cos(2\lambda_j) \right] \left[ \tilde{g}_{xkt}x_{kt}(\lambda_j) - g_{xkt}x_{kt}(\lambda_j) \right]$$

where $\tau_1$ is the first partial autocorrelation of $x_t$, which coincides with its first regular autocorrelation $\rho_1$ under the null, $x_{kt}$ is the smoothed estimate of $x_t$ based on the past, present and future values of the observed variables, $y_t$, $g_{xkt}x_{kt}(\lambda)$ its theoretical spectral density and $\tilde{g}_{xkt}x_{kt}(\lambda)$ its periodogram.

Once again, this score is entirely analogous to the univariate frequency domain score obtained if we treated $x_{kt}$ as if it was observed.

In addition, it is asymptotically orthogonal to the scores of all the parameters estimated under the null, including $\tau_1$ and $\rho_i^*$. 
ARCH(1) tests in common and specific factors

- Let us now consider tests of first order ARCH effects in the common and idiosyncratic factors.

- Our model under the alternative is:

$$y_t = \pi + cf_t + v_t,$$

$$\begin{pmatrix} f_t \\ v_t \end{pmatrix} | I_{t-1}; \theta \sim N \left[ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \lambda_t(\theta) & 0 \\ 0 & \Gamma_t(\theta) \end{pmatrix} \right],$$

$$V(f_t|I_{t-1}; \theta) = \lambda_t(\theta) = 1 + \alpha [E(f_{t-1}^2|Y_{t-1}; \theta) - 1],$$

$$V(v_{it}|I_{t-1}; \theta) = \gamma_{it}(\theta) = \gamma_i + \alpha_i^*[E(v_{it-1}^2|Y_{t-1}; \theta) - \gamma_i], \quad (i = 1, \ldots, N)$$

where \( E(f_{t-1}^2|Y_{t-j}; \theta) \) and \( E(v_{it-1}^2|Y_{t-1}; \theta) \) are the conditionally linear Kalman filter estimators of the squares of the underlying common and idiosyncratic factors obtained from this model.

- It reduces to our baseline specification under the null hypothesis that \( H_0 : \alpha^\dagger = 0 \), where \( \alpha^\dagger = (\alpha, \alpha^*) \) and \( \alpha^* = (\alpha_1, \ldots, \alpha_N) \).

- This model captures the stylised fact that correlations increase in turbulent periods.
**Proposition 5**

Let \( \bar{S}_{f_k f_k}(j) = \frac{1}{T} \sum_{t=1}^{T} \left[ f_{kt}^2(\theta_s) + \omega_k(\theta_s) - 1 \right] \left[ f_{kt-j}^2(\theta_s) + \omega_k(\theta_s) - 1 \right] \)

\[
\bar{S}_{v_k v_k}(j) = \frac{1}{T} \sum_{t=1}^{T} \text{vecd}[v_{kt}(\theta_s)v_{kt}'(\theta_s) + cc'\omega_k(\theta_s) - \Gamma] \\
\times \text{vecd}[v_{kt-j}(\theta_s)v_{kt-j}'(\theta_s) + cc'\omega_k(\theta_s) - \Gamma]
\]

denote the \( j^{th} \) sample autocovariances of the squares of the Kalman filter estimators of the innovations in the common and specific factors.

1. **Under the null hypothesis** \( H_0 : \alpha^\dagger = 0 \)

\[
LM_{ARCH(1)} = T \cdot \left( \bar{S}_{f_k f_k}(1), \text{vecd}'[\bar{S}_{v_k v_k}(1)]\Gamma^{-2} \right) \\
\times \mathcal{I}_{\alpha^\dagger \alpha^\dagger}(\theta_{s0}) \left( \bar{S}_{f_k f_k}(1), \text{vecd}'[\bar{S}_{v_k v_k}(1)]\Gamma^{-2} \right)'
\]

converges to a \( \chi^2 \) with \( N + 1 \) degrees of freedom as \( T \to \infty \), where \( \mathcal{I}_{\alpha^\dagger \alpha^\dagger}(\theta_{s0}) \) is the relevant block of the information matrix.

2. **This asymptotic null distribution is unaffected if we replace** \( \theta_{s0} \) **by its Gaussian maximum likelihood estimator** \( \bar{\theta}_s \).
ARCH(1) tests in common and specific factors

- We can re-write this score test as a joint moment test based on the $N + 1$ orthogonality conditions that we would use to test for first order ARCH effects if we treated $f_{kt}(\theta_s)$ or $v_{kit}(\theta_s)$ as the series of interest:

$$E \left\{ [f_{kt}^2(\theta_s) + \omega_k(\theta_s) - 1][f_{kt-1}^2(\theta_s) + \omega_k(\theta_s) - 1] | \theta_s, 0 \right\} = 0,$$

$$E \left\{ \gamma_i^{-2}[v_{kt}^2(\theta_s) + c_i^2 \omega_k(\theta_s) - \gamma_i][v_{kt-1}^2(\theta_s) + c_i^2 \omega_k(\theta_s) - \gamma_i] | \theta_s, 0 \right\} = 0.$$

- These moment conditions closely resemble the orthogonality conditions that we would use to test for first order ARCH effects if we could observe all the latent variables:

$$E[(f_t^2 - 1)(f_{t-1}^2 - 1) | \theta_s, 0] = 0,$$

$$E[\gamma_i^{-2}(v_{it}^2 - \gamma_i)(v_{it-1}^2 - \gamma_i) | \theta_s, 0] = 0 \quad (i = 1, \ldots, N).$$

- Our joint tests, though, take into account the non-diagonal and singular nature of the covariance matrix of $f_{kt}(\theta_s)$ and $v_{kt}(\theta_s)$, even though the latent variables themselves are uncorrelated.
Other tests for multivariate ARCH effects

- We can compare our test of ARCH effects in common and specific factors to a Hosking (1981)-type multivariate autocorrelation test applied to \( vech[(y_t - \pi)(y_t - \pi)'] \).

- Such a test would be based on

\[
\bar{S}_{yy}(j) = \frac{1}{T} \sum_{t=1}^{T} vech[(y_t - \bar{\pi})(y_t - \bar{\pi})' - \bar{\Sigma}] \\
\times vech'[(y_{t-j} - \bar{\pi})(y_{t-j} - \bar{\pi})' - \bar{\Sigma}],
\]

where \( \bar{\pi} \) and \( \bar{\Sigma} \) are the sample mean and covariance matrix of \( y_t \).

- Under the null the corresponding test statistic will be distributed as a \( \chi^2 \) with \( N^2(N+1)^2/4 \) degrees of freedom as \( T \) goes to infinity.

- The number of degrees of freedom is three orders of magnitude larger than in our test because it does not exploit the strong cross-sectional dependence of returns, which reduces power.

- It also requires \( T \) much larger than \( N^4 \) for the asymptotic distribution to be reliable in finite samples.
Power comparison: $ARCH$ tests

- We compare the power of our LM tests, Hosking’s original test, two diagonal versions, and a standard univariate first-order $ARCH$ test applied to the Equally Weighted Portfolio (EWP).

- To do so, we consider another non-exchangeable single factor model of the form:

\[
\begin{align*}
y_{it} &= \pi_i + c_i f_t + v_{it} \\
\lambda_t &= (1 - \alpha) + \alpha E(f_{t-1}^2 | Y_{t-1}) \\
\gamma_{it} &= \gamma_i (1 - \alpha^*_i) + \alpha^*_i E(v_{it-1}^2 | Y_{t-1})
\end{align*}
\]

- $\pi = (.5, .4, .5, .4, .5)$,
- $c = (5, 4, 5, 4, 5)$,
- $\gamma \propto (5, 9, 5, 9, 5)$,
- $\alpha^*_i = \alpha^* \ \forall i$.

- We evaluate power against compatible sequences of local alternatives of the form $\alpha^+_0 = \alpha^*_0 / \sqrt{T}$. 

Power comparison: $\alpha^* = \alpha$, baseline signal to noise ratio

Power of variance dependence tests versus local alternatives (common & specific)

- Joint LM
- LM on common factor
- LM on Specific factors
- Hosking
- Diagonal Hosking (vech)
- Diagonal Hosking (vecd)
- EWP

$T^{1/2} \alpha$ vs. Power
Extensions

- Models with unobserved volatility in which the conditional variances of the latent factors are functions of their own lag values, which turn out to be locally equivalent alternatives, and therefore lead to numerically identical statistics.

- \textit{ARCH}(q) tests, with or without distributed-lag or panel data type restrictions on the \textit{ARCH} coefficients.

- \textit{GARCH} tests that deal with the lack of identifiability under the null by using a Riskmetrics-type approach.

- Models with multiple factors, in which case our tests are numerically invariant to rotations of the common factors.

- Joint tests of lack of serial correlation and conditional homoskedasticity.
We can extend many of our previous results to the case in which the conditional distribution of the observed series is elliptical, but not necessarily normal.

The corresponding density will be characterised by some additional \( r \) parameters \( \eta \) that determine the shape of the conditional density of 
\[
s_t = \varepsilon_t^*/\varepsilon_t^*, \quad \text{where} \quad \varepsilon_t^* = \Sigma_t^{-1/2}(\theta_0)[y_t - \mu_t(\theta_0)].
\]

Apart from the normal distribution, which we assume corresponds to \( \eta = 0 \), we will look in some detail at a multivariate \( t \) with \( \nu_0 \) degrees of freedom.

The multivariate student \( t \) approaches the multivariate normal as \( \nu_0 \to \infty \), but has generally fatter tails.

For that reason, we shall define \( \eta \) as \( 1/\nu \), which will always remain in the finite range \( 0 \leq \eta_0 < 1/2 \) under our assumptions.
Tests that exploit leptokurtosis

- We can use the scores obtained under the assumption of ellipticity to test for serial dependence.

- The optimal score tests will be based on the previously discussed moment conditions modified by a damping factor that reflects the leptokurtosis of the distribution.

- We provide closed form expressions for the relevant blocks of the information matrix.

- We also show that the asymptotic distributions remain valid even if we estimate the damping factor using either a feasible parametric estimator of $\theta_s$ and $\eta$ obtained after fitting a specific elliptical distribution to $y_t$ under the null, or an elliptically symmetric semiparametric estimator obtained from a nonparametric estimate of the density of $\varsigma_t(\theta_s)$.
Robust Gaussian tests

- Given the intuitive nature of the associated moment conditions, we may still want to use the LM tests derived under the maintained assumption of normality when the true distribution is non-Gaussian.

- It turns out that irrespective of the ellipticity of the true distribution, it is not necessary to robustify the Gaussian tests for serial correlation that we have derived because both the Hessian matrix and the variance of the score are block diagonal with respect to $\pi$, $\rho^\dagger$ and $(c, \gamma)$, with identical blocks for $\rho^\dagger$.

- This result mimics the fact that under conditional homoskedasticity, standard score tests for serial correlation in observed series are robust to non-normality in the conditional distribution.

- In fact, since $V[f_{kt}(\theta_s)|\theta_s, 0, \eta] = c'\Sigma^{-1}(\theta_s)c$, we can obtain an asymptotically equivalent test of $H_0: \rho = 0$ by computing the $F$ test of the regression of $f_{kt}(\theta_s)$ on a constant and $f_{kt-1}(\theta_s)$, whose null distribution does not depend on Gaussianity.
Robust Gaussian tests

- In contrast, it is necessary to use sandwich expressions for robustifying the tests for ARCH in common and idiosyncratic factors.
- Luckily, the Hessian and covariance matrix of the score remain block diagonal between $\pi$, $(c, \gamma)$ and $\alpha^\dagger$, but with different blocks for $\alpha^\dagger$.
- This robust version can be regarded as the factor analytic analogue to Koenker’s (1981) suggestion for robustifying Gaussian tests of conditional homoskedasticity in univariate contexts.
- Since $V[f_{k\ell t}(\theta_s)|\theta_s, 0, \eta] = c'\Sigma^{-1}(\theta_s)c$, we can obtain an asymptotically equivalent test of $H_0: \alpha = 0$ by computing the $F$ test of the regression of $f_{k\ell t}^2(\theta_s)$ on a constant and $f_{k(t-1)}^2(\theta_s)$, whose null distribution remains valid irrespective of the normality of $f_{k\ell t}(\theta_s)$.
- But if we impose that the residual variance is $2[c'\Sigma^{-1}(\theta_s)c]^2$ instead, which is its value under normality, then the $F$ test will be incorrectly sized when the conditional distribution is not Gaussian.
- A robust joint test of $H_0: \rho^\dagger = \alpha^\dagger = 0$ can still be computed as the sum of the two separate components under the maintained assumption of ellipticity, but not under more general circumstances.
Data

- Our database consists of monthly returns on five portfolios of US stocks grouped by industry in excess of the one-month Treasury bill rate from January 1952 to December 2008.
- 672 observations as we reserve 1952 for pre-sample values.
- Industry definitions:
  1. Cnsmr: Consumer Durables, NonDurables, Wholesale, Retail, and Some Services (Laundries, Repair Shops).
  4. Hlth: Healthcare, Medical Equipment, and Drugs.
  5. Other: Other – Mines, Constr, BldMt, Trans, Hotels, Bus Serv, Entertainment, Finance.
Parameter estimates

<table>
<thead>
<tr>
<th>Sector</th>
<th>Means</th>
<th>Std.dev.</th>
<th>Cnsmr</th>
<th>Manuf</th>
<th>HiTec</th>
<th>Hlth</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cnsmr</td>
<td>.566</td>
<td>4.481</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manuf</td>
<td>.543</td>
<td>4.178</td>
<td>.804</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HiTec</td>
<td>.497</td>
<td>5.320</td>
<td>.734</td>
<td>.718</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hlth</td>
<td>.733</td>
<td>4.995</td>
<td>.710</td>
<td>.668</td>
<td>.634</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>.500</td>
<td>4.998</td>
<td>.878</td>
<td>.848</td>
<td>.739</td>
<td>.708</td>
<td>1</td>
</tr>
</tbody>
</table>

- Fraction of variances and covariances explained by single factor: 99.47
- Multivariate skewness test (Gaussian): 7.01 (p=.22)
- Multivariate kurtosis test: 1478.9 (1 df)
- Student t tail parameter: \( \hat{\eta} = .189 \) (\( \hat{\nu} = 5.3 \))
- Multivariate skewness test (Student t): 3.83 (p=.57)
# Serial correlation tests

<table>
<thead>
<tr>
<th></th>
<th>\text{PML}</th>
<th>\text{ML}</th>
<th>\text{SSP}</th>
<th>\text{PML}</th>
<th>\text{ML}</th>
<th>\text{SSP}</th>
<th>\text{PML}</th>
<th>\text{ML}</th>
<th>\text{SSP}</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Common factor</strong></td>
<td>0.35</td>
<td>2.64</td>
<td>1.35</td>
<td>19.75</td>
<td>35.49</td>
<td>24.04</td>
<td>39.59</td>
<td>53.85</td>
<td>59.63</td>
</tr>
<tr>
<td><strong>Specific factors</strong></td>
<td>1.46</td>
<td>2.70</td>
<td>1.45</td>
<td>1.40</td>
<td>8.84</td>
<td>4.11</td>
<td>0.06</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td><strong>Joint</strong></td>
<td>0.11</td>
<td>0.87</td>
<td>0.30</td>
<td>1.52</td>
<td>11.31</td>
<td>4.71</td>
<td>0.11</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

**p-values (%)**
### Conditional homoskedasticity tests

<table>
<thead>
<tr>
<th></th>
<th><strong>ARCH(1)</strong></th>
<th></th>
<th><strong>GARCH(1,1)</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PML</td>
<td>ML</td>
<td>SSP</td>
<td>PML</td>
</tr>
<tr>
<td>Common factor</td>
<td>0.36</td>
<td>6.12</td>
<td>1.79</td>
<td>0.00</td>
</tr>
<tr>
<td>Specific factors</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Joint</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

p-values (%)
Contributions of this paper

- We derive score tests of serial correlation in the levels and squares of common and idiosyncratic factors in static factor models.
- We also obtain tests against higher order serial correlation in dynamic factor models using frequency domain methods.
- We show that the implicit orthogonality conditions resemble the orthogonality conditions of models with observed factors but the weighting matrices reflect their unobservability.
- We compare our tests to existing tests for serial dependence in multivariate time series, which suffer from the curse of dimensionality.
- We robustify our Gaussian tests against non-normality, and derive more powerful versions when the conditional distribution is elliptically symmetric, which can be either parametrically or semiparametrically specified.
- We conduct Monte Carlo exercises to study the finite sample reliability and power of our proposed tests.
Contributions of this paper

- We confirm that substantial power gains accrue by exploiting the strong cross-sectional dependence of returns.
- We also show that there are additional power gains from exploiting the leptokurtosis of financial returns, as well as the persistent behaviour of conditional variances.
- We apply our methods to monthly stock returns on five US broad industry portfolios.
- We find clear evidence in favour of first order serial correlation in common and specific factors, weaker evidence for persistent components in the idiosyncratic terms, and no evidence that such a component appears in the common factor.
- We also find strong evidence for persistent serial correlation in the volatility of common and specific terms.