DYNAMIC SPECIFICATION TESTS FOR DYNAMIC FACTOR MODELS

Gabriele Fiorentini and Enrique Sentana

CEMFI Working Paper No. 1306

June 2013

CEMFI
Casado del Alisal 5; 28014 Madrid
Tel. (34) 914 290 551 Fax (34) 914 291 056
Internet: www.cemfi.es

This paper is a substantially revised version of Fiorentini and Sentana (2009), which dealt with static factor models only. We are grateful to Dante Amengual and Andrew Harvey, as well as to seminar audiences at Bologna, Columbia, ECARES/ULB, Georgetown, Penn, Princeton, Salento, Toulouse, the Oxford-Man Institute Time Series Econometrics Conference in honour of Andrew Harvey, the fifth ICEEE Congress (Genova), the first Barcelona GSE Summer Forum and the RCEA Conference (Toronto) for helpful comments, discussions and suggestions. We are also grateful to Máximo Camacho and Gabriel Pérez-Quirós for allowing us to use their data and assisting us in reproducing their results. The remarks of an associate editor and two anonymous referees also led to a radical overhaul of the paper. Of course, the usual caveat applies. Financial support from MIUR through the project .Multivariate statistical models for risk assessment.(Fiorentini) and the Spanish Ministry of Science and Innovation through grant ECO 2011-26342 (Sentana) is gratefully acknowledged.

CEMFI Working Paper 1306
June 2013

DYNAMIC SPECIFICATION TESTS FOR DYNAMIC FACTOR MODELS

Abstract

We derive computationally simple and intuitive expressions for score tests of neglected serial correlation in common and idiosyncratic factors in dynamic factor models using frequency domain techniques. The implied time domain orthogonality conditions are analogous to the conditions obtained by treating the smoothed estimators of the innovations in the latent factors as if they were observed, but they account for their final estimation errors. Monte Carlo exercises confirm the finite sample reliability and power of our proposed tests. Finally, we illustrate their empirical usefulness in an application that constructs a monthly coincident indicator for the US from four macro series.

JEL Codes: C32, C38, C52, C12, C13.
Keywords: Kalman filter, LM tests, Spectral maximum likelihood, Wiener-Kolmogorov filter.

Gabriele Fiorentini
Università di Firenze
fiorentini@disia.unifi.it

Enrique Sentana
CEMFI
sentana@cemfi.es

1 Introduction

Dynamic factor models have been extensively used in macroeconomics and finance since their introduction by Sargent and Sims (1977) and Geweke (1977) as a way of capturing the cross-sectional and dynamic correlations between multiple series in a parsimonious way. A far from comprehensive list of early and more recent applications include not only business cycle analysis (see Litterman and Sargent (1979), Stock and Watson (1989, 1991, 1993), Diebold and Rudebusch (1996) or Gregory, Head and Raynauld (1997)) and bond yields (Singleton (1981), Jegadeesh and Pennacchi (1996), Dungey, Martin and Pagan (2000) or Diebold, Rudebusch and Aruoba (2006)), but also wages (Engle and Watson (1981)), employment (Quah and Sargent (1993)), commodity prices (Peña and Box (1987)) and financial contagion (Mody and Taylor (2007)).

The model parameters are typically estimated by maximising the likelihood function of the observed data, which can be readily obtained either as a by-product of the Kalman filter prediction equations or from Whittle's (1962) frequency domain asymptotic approximation.[1] Once the parameters have been estimated, filtered values of the latent factors can be extracted by means of the Kalman smoother or its Wiener-Kolmogorov counterpart. These estimation and filtering issues are well understood (see e.g. Harvey (1989)), and the same can be said of their efficient numerical implementation (see Jungbacker and Koopman (2008)). However, several important modelling issues arise in practice, such as the right number of factors or the identification of their effects.

Another non-trivial empirical issue is the specification of the dynamics of common and idiosyncratic factors. When the cross-sectional dimension, N, is very large, one might expect to accurately recover the latent factors using simpler procedures (see Bai and Ng (2008) and the references therein). But in models in which N is small, the filtered estimates of the state variables are likely to be heavily influenced by the dynamic specification of the model, which thus becomes a first order issue. The objective of our paper is precisely to provide diagnostics for neglected serial correlation in those state variables. For that reason, we focus on Lagrange Multiplier (LM) tests, which only require estimation of the model under the null. As is well known, Likelihood ratio (LR), Wald and LM tests are equivalent under the null and sequences of local alternatives as the number of observations increases for a fixed cross-sectional dimension, and therefore they share their optimality properties.[2] In addition to computational considerations,

[1] Watson and Engle (1983) and Quah and Sargent (1993) discuss the application of the EM algorithm of Dempster, Laird and Rubin (1977) in this context, which avoids the computation of the likelihood function. As is well known, though, this algorithm slows down considerably near the optimum, so it is best used as a procedure for obtaining good initial values.

[2] Extensions to situations in which both data dimensions simultaneously grow are left for further research.

which are particularly relevant when one is concerned about several alternatives, an important advantage of LM tests expressed as score tests is that they often coincide with tests of easy to interpret moment conditions (see Newey (1985) and Tauchen (1985)), which will continue to have non-trivial power even in situations for which they are not optimal. As we shall see, our proposed tests are no exception in that regard.

Earlier work on specification testing in dynamic factor models include Engle and Watson (1980), who explained how to apply the LM testing principle in the time domain for models with static factor loadings, Geweke and Singleton (1981), who studied LR and Wald tests in the frequency domain, and Fernández (1990), who applied the LM principle in the frequency domain to a multivariate "structural time series model" (see Harvey (1989) for a comparison of time domain and frequency domain testing methods in that context).

Aside from considering a general class of models, our main contribution is that our proposed tests are very simple to implement, and even simpler to interpret. Once a model has been specified and estimated, score tests focusing on several departures from the null can be routinely computed from simple statistics of the estimated state variables. And even though our theoretical derivations make extensive use of spectral methods for time series, we provide both time domain and frequency domain interpretations of the relevant scores, so researchers who strongly prefer one method over the other could apply them without abandoning their favourite estimation techniques.

The rest of the paper is organised as follows. In section 2, we review the properties of dynamic factor models, their estimators and filters. Then, we derive our tests in section 3, and present a Monte Carlo evaluation of their finite sample behaviour in section 4. This is followed in section 5 by an empirical illustration that revisits the dynamic factor model used by Camacho, Pérez-Quirós and Poncela (2012) to construct a coincident indicator for the US. Finally, our conclusions, together with several interesting extensions, can be found in section 6. Auxiliary results are gathered in appendices.

2 Theoretical background

2.1 Dynamic factor models

To keep the notation to a minimum, we focus on single factor models, which suffice to illustrate our main results. A parametric version of a dynamic exact factor model for a finite dimensional vector of N observed series, $\mathbf{y}_t$, can be defined in the time domain by the system of equations

$$
\begin{aligned}
\mathbf{y}_t &= \boldsymbol{\pi} + \mathbf{c}(L)x_t + \mathbf{u}_t, \\
\alpha_x(L)x_t &= \beta_x(L)f_t, \\
\alpha_{u_i}(L)u_{i,t} &= \beta_{u_i}(L)v_{i,t}, \quad i = 1, \ldots, N, \\
(f_t, v_{1,t}, \ldots, v_{N,t})|I_{t-1}; \boldsymbol{\pi}, \boldsymbol{\theta} &\sim N[0, diag(1, \gamma_1, \ldots, \gamma_N)],
\end{aligned}
$$

where x_t is the common factor, $\mathbf{u}_t$ the N specific factors, $\mathbf{c}(L) = \sum_{\ell=-F}^{M} \mathbf{c}_\ell L^\ell$ a vector of possibly two-sided polynomials in the lag operator, $\alpha_x(L)$ and $\alpha_{u_i}(L)$ are one-sided polynomials of orders p_x and p_{u_i}, respectively, while $\beta_x(L)$ and $\beta_{u_i}(L)$ are one-sided (coprime) polynomials of orders q_x and q_{u_i}, I_{t-1} is an information set that contains the values of $\mathbf{y}_t$ and f_t up to, and including time $t-1$, $\boldsymbol{\pi}$ is the mean vector and $\boldsymbol{\theta}$ refers to all the remaining model parameters.[3]

A specific example would be

$$
\begin{aligned}
\begin{pmatrix} y_{1,t} \\ \vdots \\ y_{N,t} \end{pmatrix} &= \begin{pmatrix} \pi_1 \\ \vdots \\ \pi_N \end{pmatrix} + \begin{pmatrix} c_{1,0} \\ \vdots \\ c_{N,0} \end{pmatrix} x_t + \begin{pmatrix} c_{1,1} \\ \vdots \\ c_{N,1} \end{pmatrix} x_{t-1} + \begin{pmatrix} u_{1,t} \\ \vdots \\ u_{N,t} \end{pmatrix}, \\
x_t &= \alpha_{x1}x_{t-1} + f_t, \\
u_{it} &= \alpha_{u_i1}u_{it-1} + v_{it}, \quad i = 1, \ldots, N.
\end{aligned} \tag{1}
$$

Note that the dynamic nature of the model is the result of three different characteristics:

1. The serial correlation of the common factor
2. The serial correlation of the idiosyncratic factors
3. The dynamic impact of the common factor on the observed variables.

Thus, we would need to shut down all three sources to go back to a traditional static factor model (see Lawley and Maxwell (1971)). Cancelling only one or two of those channels still results in a dynamic factor model. For example, Engle and Watson (1981) considered models with static factor loadings, while Peña and Box (1987) further assumed that the specific factors were white noise. To some extent, characteristics 1 and 3 overlap, as one could always write any dynamic factor model in terms of white noise common factors. In this regard, the assumption of $\textsc{Arma}(p_x, q_x)$ dynamics for the common factor can be regarded as a parsimonious way of modelling an infinite distributed lag.[4]

[3] We could relax the assumption of cross-sectional orthogonality in the idiosyncratic terms, but in general we would still need to impose some parametric restrictions for identification purposes given that we maintain the assumption of fixed N.

[4] Some dynamic factor models can be written as static factor models with a larger number of factos. For example, in model (1) we could define f_t and x_{t-1} as two "orthogonal" static factors, with factor loading $c_{i,0}$ and $c_{i,1} + \alpha_{x1}c_{i,0}$ respectively. Our tests, though, apply to all factor models, including those without a static factor representation.

In this paper we are interested in hypothesis tests for $p_x = d_x$ vs $p_x = d_x + k_x$ or $p_{u_i} = d_{u_i}$ vs $p_{u_i} = d_{u_i} + k_{u_i}$, or the analogous hypotheses for q_x and q_{u_i}. To avoid dealing with nonsensical situations, we maintain the assumption that the model which has been estimated under the null is identified (see Geweke (1977) and Geweke and Singleton (1981) for a general discussion of identification in dynamic factor models, and Heaton and Solo (2004) for more specific results for the parametric models that we consider in this paper).

2.2 Tests of white noise vs. AR(1) in the common factors

Let us start by quickly reviewing the first order serial correlation tests obtained by Fiorentini and Sentana (2012). The baseline model in that paper is the static factor model

$$\begin{gathered}
\mathbf{y}_t = \boldsymbol{\pi} + \mathbf{c}x_t + \mathbf{u}_t, \\
x_t = f_t, \\
\mathbf{u}_t = \mathbf{v}_t, \\
\begin{pmatrix} f_t \\ \mathbf{v}_t \end{pmatrix} | I_{t-1}, \boldsymbol{\theta}_s \sim N\left[\begin{pmatrix} 0 \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Gamma} \end{pmatrix}\right],
\end{gathered}$$

which remains rather popular in finance (except in term structure applications) (see Connor, Goldberg and Korajczik (2010) and the references therein).

The Kalman smoother yields the same factor estimates as the Kalman filter updating equations, which have simple closed form expressions:

$$\begin{aligned}
f_{t|t} &= f_{t|T} = \mathbf{c}'\boldsymbol{\Sigma}^{-1}(\mathbf{y}_t - \boldsymbol{\pi}) = \frac{\mathbf{c}'\boldsymbol{\Gamma}^{-1}}{1 + \mathbf{c}'\boldsymbol{\Gamma}^{-1}\mathbf{c}}(\mathbf{y}_t - \boldsymbol{\pi}), \\
\mathbf{v}_{t|t} &= \mathbf{v}_{t|T} = \boldsymbol{\Gamma}\boldsymbol{\Sigma}^{-1}(\mathbf{y}_t - \boldsymbol{\pi}) = \mathbf{y}_t - \boldsymbol{\pi} - \mathbf{c}x_{t|t}.
\end{aligned}$$

A potentially interesting alternative would be:

$$\begin{gathered}
\mathbf{y}_t = \boldsymbol{\pi} + \mathbf{c}x_t + \mathbf{u}_t, \\
x_t = \psi x_{t-1} + f_t, \\
\mathbf{u}_t = \mathbf{v}_t.
\end{gathered}$$

This alternative reduces to the static specification under the null $H_0 : \psi = 0$. Otherwise, it has the autocorrelation structure of a VARMA(1,1). Fiorentini and Sentana (2012) show that testing the null of multivariate white noise against such a complex VARMA(1,1) specification is extremely easy. Specifically, they show that the average score with respect to ψ under H_0 is

$$\bar{\mathbf{s}}_{\psi T} = \frac{1}{T}\sum_{t=2}^{T} f_{t|T} f_{t-1|T},$$

which is entirely analogous to the score that one would use to test for first order serial correlation in f_t if the latent factors were observed (see Breusch and Pagan (1980) or Godfrey (1989)). The main difference is that the asymptotic variance of this score is $[\mathbf{c}'\boldsymbol{\Sigma}^{-1}\mathbf{c}]^2 < 1$. Fiorentini and Sentana (2012) interpret $\mathbf{c}'\boldsymbol{\Sigma}^{-1}\mathbf{c}$ as the R^2 in the theoretical least squares projection of f_t on a constant and $\mathbf{y}_t$. Therefore, the higher the degree of observability of the common factor, the closer the asymptotic variance of the average score will be to 1, which is the asymptotic variance of the first sample autocorrelation of f_t. Intuitively, this convergence result simply reflects the fact that the common factor becomes observable as the "signal to noise" ratio $\mathbf{c}'\mathbf{\Sigma}^{-1}\mathbf{c}$ approaches 1. Before the limit, though, the test takes into account the unobservability of f_t. Given that $\mathbf{c}'\mathbf{\Sigma}^{-1}\mathbf{c} = (\mathbf{c}'\mathbf{\Gamma}^{-1}\mathbf{c})/[1+(\mathbf{c}'\mathbf{\Gamma}^{-1}\mathbf{c})]$ under the assumption that $\mathbf{\Gamma}$ has full rank, the aforementioned R^2 will typically be close to 1 for N large due to the pervasive nature of the common factor (see e.g. Sentana (2004)).

When we move to testing say AR(1) vs AR(2) in the unobservable factors, the model is already dynamic under the null and the Kalman filter and smoother equations no longer coincide. More importantly, those equations are recursive and therefore difficult to characterise without solving a multivariate algebraic Riccati equation. 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 (see Appendix A).

An alternative way to characterise a dynamic factor model is in the frequency domain. As we shall see, the (non-recursive) frequency domain scores remain remarkably simple, since they closely resemble the scores of a static factor model.

2.3 Maximum likelihood estimation in the frequency domain

In what follows, we maintain the assumption that $\mathbf{y}_t$ is a covariance stationary process, possibly after suitable transformations as in section 5.

Under stationarity, the spectral density matrix of the observed variables is proportional to

$$\mathbf{G_{yy}}(\lambda) = \mathbf{c}(e^{-i\lambda t})G_{xx}(\lambda)\mathbf{c}'(e^{i\lambda t}) + \mathbf{G_{uu}}(\lambda),$$
$$G_{xx}(\lambda) = \frac{\beta_x(e^{-i\lambda t})\beta_x(e^{i\lambda t})}{\alpha_x(e^{-i\lambda t})\alpha_x(e^{i\lambda t})},$$
$$\mathbf{G_{uu}}(\lambda) = diag[G_{u_1u_1}(\lambda), \ldots, G_{u_Nu_N}(\lambda)],$$
$$G_{u_iu_i}(\lambda) = \gamma_i \frac{\beta_{u_i}(e^{-i\lambda})\beta_{u_i}(e^{i\lambda t})}{\alpha_{u_i}(e^{-i\lambda t})\alpha_{u_i}(e^{i\lambda t})},$$

which inherits the exact single factor structure of the unconditional covariance matrix of a static factor model. Let

$$\mathbf{I_{yy}}(\lambda) = \frac{1}{2\pi T}\sum_{t=1}^{T}\sum_{s=1}^{T}(\mathbf{y}_t - \boldsymbol{\pi})(\mathbf{y}_s - \boldsymbol{\pi})'e^{-i(t-s)\lambda} \quad (2)$$

denote the periodogram matrix and $\lambda_j = 2\pi j/T$ $(j = 0, \ldots T-1)$ the usual Fourier frequencies. If we assume that $\mathbf{G_{yy}}(\lambda)$ is not singular at all frequencies,[5] the so-called Whittle (discrete

[5] Otherwise, there would be a linear combination of the components of the $\mathbf{y}_t's$ at frequency λ that would be identically 0.

spectral) approximation to the log-likelihood function is[6]

$$-\frac{NT}{2}\ln(2\pi)-\frac{1}{2}\sum_{j=0}^{T-1}\ln|\mathbf{G}_{\mathbf{yy}}(\lambda_j)|-\frac{1}{2}\sum_{j=0}^{T-1}\operatorname{tr}\left\{\mathbf{G}_{\mathbf{yy}}^{-1}(\lambda_j)[2\pi\mathbf{I}_{\mathbf{yy}}(\lambda_j)]\right\}. \tag{3}$$

If we further assume that $G_{xx}(\lambda)>0$ and $G_{u_iu_i}(\lambda)>0$ for all i, computations can be considerably speeded up by exploiting that

$$\mathbf{G}_{\mathbf{yy}}^{-1}(\lambda)=\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)-\omega(\lambda)\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda t})\mathbf{c}'(e^{i\lambda t})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda),$$

$$\omega(\lambda)=[G_{xx}^{-1}(\lambda)+\mathbf{c}'(e^{i\lambda t})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda t})]^{-1}.$$

The MLE of $\boldsymbol{\pi}$, which only enters through $\mathbf{I}_{\mathbf{yy}}(\lambda)$, is the sample mean, so in what follows we focus on demeaned variables. In turn, the score with respect to all the remaining parameters is

$$\mathbf{d}(\boldsymbol{\theta})=\frac{1}{2}\sum_{j=0}^{T-1}\frac{\partial vec'[\mathbf{G}_{\mathbf{yy}}(\lambda_j)]}{\partial\boldsymbol{\theta}}\mathbf{M}(\lambda_j)\mathbf{m}(\lambda_j),$$

$$\mathbf{m}(\lambda)=vec\left[2\pi\mathbf{I}_{\mathbf{yy}}'(\lambda)-\mathbf{G}_{\mathbf{yy}}'(\lambda)\right],$$

$$\mathbf{M}(\lambda)=\left[\mathbf{G}_{\mathbf{yy}}^{-1}(\lambda)\otimes\mathbf{G}_{\mathbf{yy}}^{-1\prime}(\lambda)\right].$$

We provide numerically reliable and fast to compute expressions for all the required derivatives in Appendix B.

The information matrix is

$$\mathbf{Q}=\frac{1}{4\pi}\int_{-\pi}^{\pi}\frac{\partial vec'[\mathbf{G}_{\mathbf{yy}}(\lambda)]}{\partial\boldsymbol{\theta}}\mathbf{M}(\lambda)\left\{\frac{\partial vec'[\mathbf{G}_{\mathbf{yy}}(\lambda)]}{\partial\boldsymbol{\theta}}\right\}^{*}d\lambda,$$

where * denotes the conjugate transpose of a matrix. A consistent estimator will be provided by either by the outer product of the score or by

$$\boldsymbol{\Phi}(\boldsymbol{\theta})=\frac{1}{2}\sum_{j=0}^{T-1}\frac{\partial vec'[\mathbf{G}_{\mathbf{yy}}(\lambda_j)]}{\partial\boldsymbol{\theta}}\mathbf{M}(\lambda_j)\left\{\frac{\partial vec'[\mathbf{G}_{\mathbf{yy}}(\lambda)]}{\partial\boldsymbol{\theta}}\right\}^{*}.$$

Formal results showing the strong consistency and asymptotic normality of the resulting ML estimators under suitable regularity conditions have been provided by Dunsmuir and Hannan (1976) and Dunsmuir (1979), who also show their asymptotic equivalence to the time domain ML estimators.[7]

[6] There is also a continuous version which replaces sums by integrals (see Dusmuir and Hannan (1976)).

[7] This equivalence is not surprising in view of the contiguity of the Whittle measure in the Gaussian case (see Choudhuri, Ghosal and Roy (2004)).

2.4 The (Kalman-)Wiener-Kolmogorov filter

By working in the frequency domain we can easily obtain smoothed estimators of the latent variables too. Specifically, let

$$\begin{aligned} \mathbf{y}_t - \boldsymbol{\pi} &= \int_{-\pi}^{\pi} e^{i\lambda t} d\mathbf{Z}_{\mathbf{y}}(\lambda), \\ V[d\mathbf{Z}_{\mathbf{y}}(\lambda)] &= \mathbf{G}_{\mathbf{yy}}(\lambda) d\lambda \end{aligned}$$

denote Cramer's spectral decomposition of the observed process, which is the frequency domain analogue to Wold's decomposition.

The Wiener-Kolmogorov two-sided filter for the common factor x_t at each frequency is given by

$$\mathbf{c}'(e^{i\lambda})G_{xx}(\lambda)\mathbf{G}_{\mathbf{yy}}^{-1}(\lambda)d\mathbf{Z}_{\mathbf{y}}(\lambda)$$

so that the spectral density of the smoother $x_{t|T}^K$ as $T \to \infty$[8] will be

$$G_{x^K x^K}(\lambda) = \mathbf{c}'(e^{i\lambda})G_{xx}(\lambda)\mathbf{G}_{\mathbf{yy}}^{-1}(\lambda)G_{xx}(\lambda)\mathbf{c}(e^{-i\lambda}) = \omega(\lambda)\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda})G_{xx}(\lambda). \quad (4)$$

Hence, the spectral density of the final estimation error $x_t - x_{t|\infty}^K$ will be given by

$$G_{xx}(\lambda) - \mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{yy}}^{-1}(\lambda_j)\mathbf{c}(e^{i\lambda}) = \omega(\lambda).$$

Having obtained these, we can easily obtain the smoother for f_t, $f_{t|\infty}^K$, by applying to $x_{t|\infty}^K$ the one-sided filter

$$\alpha_x(e^{-i\lambda})/\beta_x(e^{-i\lambda})$$

Likewise, we can derive its spectral density, as well as the spectral density of its final estimation error $f_t - f_{t|\infty}^K$. Finally, we can obtain the autocovariances of $x_{t|\infty}^K$, $f_{t|\infty}^K$ and their final estimation errors by applying the usual inverse Fourier transformation

$$cov(z_t, z_{t-k}) = \int_{-\pi}^{\pi} e^{i\lambda k} G_{zz}(\lambda) d\lambda.$$

2.5 The minimal sufficient statistics for $\{x_t\}$

In any given realisation of the vector process $\{\mathbf{y}_t\}$, the values of $\{x_t\}$ could be regarded as a set of T parameters. With this interpretation in mind, we can define $x_{t|\infty}^G$ as the spectral GLS estimator of x_t through the transformation

$$[\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda})]^{-1}\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)d\mathbf{Z}_{\mathbf{y}}(\lambda).$$

[8] The main difference between the Wiener-Kolmogorov filtered values, $x_{t|\infty}^K$, and the Kalman filter smoothed values, $x_{t|T}^K$, results from the dependence of the former on a double infinite sequence of observations. As shown by Fiorentini (1995) and Gómez (1999), though, they can be made numerically identical by replacing both pre- and post- sample observations by their least squares projections onto the linear span of the sample observations.

Similarly, we can define $\mathbf{u}^G_{t|\infty}$ though

$$\{\mathbf{I}_N - \mathbf{c}(e^{-i\lambda})[\mathbf{c}'(e^{i\lambda})\mathbf{G}^{-1}_{\mathbf{uu}}(\lambda)\mathbf{c}(e^{-i\lambda})]^{-1}\mathbf{c}'(e^{i\lambda})\mathbf{G}^{-1}_{\mathbf{uu}}(\lambda)\}d\mathbf{Z}_{\mathbf{y}}(\lambda).$$

It is then easy to see that the joint spectral density of $x^G_{t|\infty}$ and $\mathbf{u}^G_{t|\infty}$ will be block-diagonal, with the (1,1) element being

$$G_{xx}(\lambda) + [\mathbf{c}'(e^{i\lambda})\mathbf{G}^{-1}_{\mathbf{uu}}(\lambda)\mathbf{c}(e^{-i\lambda})]^{-1}$$

and the (2,2) block

$$\mathbf{G}_{\mathbf{yy}}(\lambda) - \mathbf{c}(e^{-i\lambda})[\mathbf{c}'(e^{i\lambda})\mathbf{G}^{-1}_{\mathbf{uu}}(\lambda)\mathbf{c}(e^{-i\lambda})]^{-1}\mathbf{c}'(e^{i\lambda}),$$

whose rank is $N-1$. This orthogonalisation may be regarded as the frequency domain version of the endogenous factorial representation in Gourieroux, Monfort and Renault (1991). As such, it allows us to factorise the spectral log-likelihood function of $\mathbf{y}_t$ as the sum of the log-likelihood function of $x^G_{t|\infty}$, which is univariate, and the log-likelihood function of $\mathbf{u}^G_{t|\infty}$.[9] Importantly, the parameters characterising $G_{xx}(\lambda)$ only enter through the first component. In contrast, the remaining parameters affect both components. Moreover, we can easily show that

1. $x^G_{t|\infty} = x_t + \zeta^G_{t|\infty}$, with x_t and $\zeta^G_{t|\infty}$ orthogonal at all leads and lags[10]
2. The smoothed estimator of x_t obtained by applying the Wiener- Kolmogorov filter to $x^G_{t|\infty}$ coincides with $x^K_{t|\infty}$.

This confirms that $x^G_{t|\infty}$ constitute minimal sufficient statistics for x_t, thereby generalising earlier results by Fiorentini, Sentana and Shephard (2004), who looked at the related class of factor models with time-varying volatility, and Jungbacker and Koopman (2008), who considered models in which $\mathbf{c}(e^{-i\lambda}) = \mathbf{c}$ for all λ.[11]

[9] The Jacobian of the transformation is 1, as we can write

$$\begin{pmatrix} [\mathbf{c}'(e^{i\lambda})\mathbf{G}^{-1}_{\mathbf{uu}}(\lambda)\mathbf{c}(e^{-i\lambda})]^{-1}\mathbf{c}'(e^{i\lambda})\mathbf{G}^{-1}_{\mathbf{uu}}(\lambda) \\ \{\mathbf{I}_N - \mathbf{c}(e^{-i\lambda})[\mathbf{c}'(e^{i\lambda})\mathbf{G}^{-1}_{\mathbf{uu}}(\lambda)\mathbf{c}(e^{-i\lambda})]^{-1}\mathbf{c}'(e^{i\lambda})\mathbf{G}^{-1}_{\mathbf{uu}}(\lambda)\} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{G}^{1/2}_{\mathbf{uu}}(\lambda) \end{pmatrix} \begin{pmatrix} [\mathbf{c}'(e^{i\lambda})\mathbf{G}^{-1}_{\mathbf{uu}}(\lambda)\mathbf{c}(e^{-i\lambda})]^{-1}\mathbf{c}'(e^{i\lambda})\mathbf{G}^{-1/2}_{\mathbf{uu}}(\lambda) \\ \{\mathbf{I}_N - \mathbf{G}^{-1/2}_{\mathbf{uu}}(\lambda)\mathbf{c}(e^{-i\lambda})[\mathbf{c}'(e^{i\lambda})\mathbf{G}^{-1}_{\mathbf{uu}}(\lambda)\mathbf{c}(e^{-i\lambda})]^{-1}\mathbf{c}'(e^{i\lambda})\mathbf{G}^{-1/2}_{\mathbf{uu}}(\lambda)\} \end{pmatrix} \mathbf{G}^{-1/2}_{\mathbf{uu}}(\lambda),$$

where the matrix in the centre is orthogonal.

[10] This implies that $E(x^G_{t|\infty}|x_t) = x_t$, which confirms that while $x^K_{t|\infty}$ can be understood as a Bayesian cross-sectional GLS estimator of x_t that uses the prior $x_t \sim N[0, G_{xx}(\lambda)]$, $x^G_{t|\infty}$ relies on a diffuse prior instead.

[11] It is also possible to relate $x^G_{t|\infty}$ to the first spectral principal component extracted from $\mathbf{G}^{-1/2}_{\mathbf{uu}}(\lambda)\mathbf{G}_{\mathbf{yy}}(\lambda)\mathbf{G}^{-1/2}_{\mathbf{uu}}(\lambda)$ along the lines of Appendix 2 in Sentana (2004).

2.6 Autocorrelation structure of the factor estimators

As discussed in Maravall (1999), the serial dependence structure of the estimators of the latent variables can be a useful tool for model diagnostic. Large discrepancies between theoretical and empirical autocovariance functions of the factor estimators can be interpreted as indication of model misspecification. As we shall see below, our LM tests carry out this comparison in a very precise statistical sense.

Smoothed factors, though, are the result of optimal symmetric two-sided filters. As a consequence, their serial correlation structure is generally different from that of the unobserved state variables. Specifically, the frequency by frequency orthogonality of predictor and prediction error implies that $G_{x^K x^K}(\lambda) \leq G_{xx}(\lambda)$ for all λ, so that the smoothed estimates are smoother than the latent factors. In addition, the degree of unobservability of x_t depends exclusively on the size of $[\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda})]^{-1}$ relative to $G_{xx}(\lambda)$, which is generally different for different frequencies. This can be visualised by representing over $[-\pi, \pi]$ either $G_{x^K x^K}(\lambda)$, $[\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda})]^{-1}$ and their sum $G_{xx}(\lambda)$, or the R^2-type, signal to noise measure $G_{x^K x^K}(\lambda)/G_{xx}(\lambda)$.

In our general multivariate setting, the time domain structure of the smoothed components is complicated and difficult to interpret. There are special cases, however, in which the resulting models for the unobserved factors are rather simple. Consider first the case where $\beta_x(L) = \beta_{u_1}(L) = \cdots = \beta_{u_N}(L) = 1$, so that all state variables follow purely autoregressive processes. Moreover, assume static loadings to simplify the exposition. The Fourier transform of $G_{x^K x^K}(\lambda)$ yields the autocovariance generating function (Acgf) of x^K, which is given by

$$\begin{aligned} ACGF_{x^K}(L) &= \frac{\sum_{i=1}^N c_i^2 \alpha_{u_i}(L)\alpha_{u_i}(L^{-1})/\gamma_i}{\left(\sum_{i=1}^N c_i^2 \alpha_{u_i}(L)\alpha_{u_i}(L^{-1})/\gamma_i\right) + \alpha_x(L)\alpha_x(L^{-1})} \frac{1}{\alpha_x(L)\alpha_x(L^{-1})} \\ &= \frac{\beta_{x^K}(L)\beta_{x^K}(L^{-1})}{\alpha_{x^K}(L)\alpha_{x^K}(L^{-1})}\gamma_{x^K}, \end{aligned}$$

where γ_{x^K} denotes the variance of the univariate Wold innovations in $x^K_{t|\infty}$.

Let $\overline{p}_{\mathbf{u}} = \max(p_{u_i})$ and $\overline{p} = \max(\overline{p}_{\mathbf{u}}, p_x)$. Then, it is easy to prove that $\beta_{x^K}(L)$ and $\alpha_{x^K}(L)$ are polynomials of orders $\overline{p}_{\mathbf{u}}$ and $\overline{p} + p_x$ respectively. Hence, the factor estimators will display the Acgf of and Arma($\overline{p} + p_x$, $\overline{p}_{\mathbf{u}}$). For example, when both common and specific factors follow AR(1) processes the factor estimators will display the autocorrelation of an Arma(2,1).

It is also interesting to consider the special case in which the autoregressive polynomials $\alpha_x(L)$ and $\alpha_{u_i}(L)$ share some or even all their roots. In this latter case $\alpha_x(L) = \alpha_{u_i}(L) = \alpha(L)$, and

$$ACGF_{x^K}(L) = \frac{\alpha(L)\alpha(L^{-1})\sum_{i=1}^N c_i^2/\gamma_i}{\alpha(L)\alpha(L^{-1})\left(\sum_{i=1}^N c_i^2/\gamma_i + 1\right)} \frac{1}{\alpha(L)\alpha(L^{-1})} = \frac{\sum_{i=1}^N c_i^2/\gamma_i}{\sum_{i=1}^N c_i^2/\gamma_i + 1} \frac{1}{\alpha(L)\alpha(L^{-1})}.$$

In this particular case the model for the common factor estimators is exactly the same as the model for the unobserved factor re-scaled by the static signal to noise ratio

$$\frac{\sum_{i=1}^{N} c_i^2/\gamma_i}{\sum_{i=1}^{N} c_i^2/\gamma_i + 1}.$$

Moreover, the smoother of the innovations in the common factor, $f_{t|\infty}^K$, will be white noise. Interestingly,

$$G_{x^K x^K}(\lambda) = \frac{\sum_{i=1}^{N} c_i^2/\gamma_i}{\sum_{i=1}^{N} c_i^2/\gamma_i + 1} G_{xx}(\lambda),$$

so that the ratio between the smoothed and the unobservable factor spectra is constant at all frequencies. Intuitively, this is due to the fact that in this special case the observable variables follow a VAR(1) with a diagonal companion matrix whose innovations covariance matrix retains the static factor model properties. As a consequence, the quasi-differenced data will have the usual static factor structure. Another way of obtaining the same result is by noticing that the dynamic GLS factor estimators, $x_{t|\infty}^G$, will be a static transformation of the observed series in this case.

More generally, when the common and specific factors follow ARMA processes we have that

$$ACGF_{x^K}(L) = \frac{\sum_{i=1}^{N} c_i^2 \frac{\alpha_{u_i}(L)\alpha_{u_i}(L^{-1})}{\beta_{u_i}(L)\beta_{u_i}(L^{-1})}\gamma_i}{\left(\sum_{i=1}^{N} c_i^2 \frac{\alpha_{u_i}(L)\alpha_{u_i}(L^{-1})}{\beta_{u_i}(L)\beta_{u_i}(L^{-1})}\gamma_i\right) + \frac{\alpha_x(L)\alpha_x(L^{-1})}{\beta_x(L)\beta_x(L^{-1})}} \frac{\beta_x(L)\beta_x(L^{-1})}{\alpha_x(L)\alpha_x(L^{-1})} =$$

$$= \frac{\frac{\alpha_{\mathbf{u}}(L)\alpha_{\mathbf{u}}(L^{-1})}{\beta_{\mathbf{u}}(L)\beta_{\mathbf{u}}(L^{-1})}\gamma_{\mathbf{u}}}{\frac{\alpha_{\mathbf{u}}(L)\alpha_{\mathbf{u}}(L^{-1})}{\beta_{\mathbf{u}}(L)\beta_{\mathbf{u}}(L^{-1})}\gamma_{\mathbf{u}} + \frac{\alpha_x(L)\alpha_x(L^{-1})}{\beta_x(L)\beta_x(L^{-1})}} \frac{\beta_x(L)\beta_x(L^{-1})}{\alpha_x(L)\alpha_x(L^{-1})} =$$

$$= \frac{\frac{\alpha_{\mathbf{u}}(L)\alpha_{\mathbf{u}}(L^{-1})\beta_x(L)\beta_x(L^{-1})}{\beta_{\mathbf{u}}(L)\beta_{\mathbf{u}}(L^{-1})}}{\frac{[\beta_x(L)\beta_x(L^{-1})\alpha_{\mathbf{u}}(L)\alpha_{\mathbf{u}}(L^{-1})+\alpha_x(L)\alpha_x(L^{-1})\beta_{\mathbf{u}}(L)\beta_{\mathbf{u}}(L^{-1})]\alpha_x(L)\alpha_x(L^{-1})}{\beta_{\mathbf{u}}(L)\beta_{\mathbf{u}}(L^{-1})\beta_x(L)\beta_x(L^{-1})}}\gamma_{x^K} =$$

$$\frac{\alpha_{\mathbf{u}}(L)\alpha_{\mathbf{u}}(L^{-1})\beta_x(L)\beta_x(L^{-1})\beta_x(L)\beta_x(L^{-1})}{[\beta_x(L)\beta_x(L^{-1})\alpha_{\mathbf{u}}(L)\alpha_{\mathbf{u}}(L^{-1}) + \alpha_x(L)\alpha_x(L^{-1})\beta_{\mathbf{u}}(L)\beta_{\mathbf{u}}(L^{-1})]\,\alpha_x(L)\alpha_x(L^{-1})}\gamma_{x^K}.$$

If we further assume that the moving average polynomials of the specific factor processes are coprime, then

$$\alpha_{\mathbf{u}}(L) = \sum_{i=1}^{N} c_i \alpha_{u_i}(L)\beta_{u_{\setminus i}}(L),$$

with

$$\beta_{u_{\setminus i}}(L) = \prod_{\substack{j=1 \\ j\neq i}} \beta_{u_j}(L)$$

and

$$\beta_{\mathbf{u}}(L) = \prod_{i=1} \beta_{u_i}(L).$$

For example, if all the factors are ARMA(1,1) then the order of both $\alpha_{\mathbf{u}}(L)$ and $\beta_{\mathbf{u}}(L)$ will be N, so the ACGF of the factor estimators is that of an ARMA$(N+2, N+2)$.

Let us now turn to the specific factors. The spectral matrix of the idiosyncratic smoother is

$$\mathbf{G}_{\mathbf{u}^K\mathbf{u}^K}(\lambda) = \mathbf{G}_{\mathbf{uu}}(\lambda)\mathbf{G}_{\mathbf{yy}}(\lambda)^{-1}\mathbf{G}_{\mathbf{uu}}(\lambda) = \mathbf{G}_{\mathbf{uu}}(\lambda) - \frac{\mathbf{cc}'}{\mathbf{c}'\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}+G_{xx}(\lambda)^{-1}}.$$

Similarly, the relation between the smoother and the unobserved factor spectral matrix is

$$\mathbf{G}_{\mathbf{u}^K\mathbf{u}^K}(\lambda) = \left[\mathbf{I} - \frac{\mathbf{cc}'}{\mathbf{c}'\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}+G_{xx}(\lambda)^{-1}}\right]\mathbf{G}_{\mathbf{uu}}(\lambda).$$

The Fourier transform of $\mathbf{G}_{\mathbf{u}^K\mathbf{u}^K}(\lambda)$ yields the ACGF of the VARMA process for $\mathbf{u}_t^K$, which is generally rather complicated. In the case of purely autoregressive unobserved factors, the generic i^{th} element of $vecd\,[\mathbf{G}_{\mathbf{u}^K\mathbf{u}^K}(\lambda)]$ has the form

$$\begin{aligned}
vecd\,[\mathbf{G}_{\mathbf{u}^K\mathbf{u}^K}(\lambda)]_i &= \frac{\gamma_i}{\alpha_{u_i}(e^{-i\lambda})\alpha_{u_i}(e^{i\lambda})} - \frac{c_i^2}{\sum_{i=1}^N c_i^2\alpha_{u_i}(e^{-i\lambda})\alpha_{u_i}(e^{i\lambda})/\gamma_i + \alpha_x(e^{-i\lambda})\alpha_x(e^{i\lambda})} = \\
&= \frac{\gamma_i\alpha_x(e^{-i\lambda})\alpha_x(e^{i\lambda}) + \sum_{i=1}^N c_i^2\alpha_{u_i}(e^{-i\lambda})\alpha_{u_i}(e^{i\lambda}) - c_i^2\alpha_{u_i}(e^{-i\lambda})\alpha_{u_i}(e^{i\lambda})}{\alpha_{u_i}(e^{-i\lambda})\alpha_{u_i}(e^{i\lambda})\left[\sum_{i=1}^N c_i^2\alpha_{u_i}(e^{-i\lambda})\alpha_{u_i}(e^{i\lambda})/\gamma_i + \alpha_x(e^{-i\lambda})\alpha_x(e^{i\lambda})\right]} = \\
&= \frac{\sum_{\substack{j=1\\ j\neq i}}^N c_i^2\alpha_{u_i}(e^{-i\lambda})\alpha_{u_i}(e^{i\lambda})}{\alpha_{u_i}(e^{-i\lambda})\alpha_{u_i}(e^{i\lambda})\left[\sum_{i=1}^N c_i^2\alpha_{u_i}(e^{-i\lambda})\alpha_{u_i}(e^{i\lambda})/\gamma_i + \alpha_x(e^{-i\lambda})\alpha_x(e^{i\lambda})\right]}.
\end{aligned}$$

Hence, if we call $\overline{p}_{\mathbf{u}/i} = \max_{j\neq i}(p_{u_j}, p_x)$, it is easy to see that $u_{i,t}^K$ displays the autocorrelation structure of an ARMA$(\overline{p}+p_i, \overline{p}_{\mathbf{u}/i})$.

2.7 Testing AR(1) vs AR(2) for observable x_t

Although all previous spectral calculations are straightforward, they might seem daunting unless one is familiar with frequency domain methods. Fortunately, they have remarkably simple time domain counterparts. For pedagogical purposes, let us initially assume that x_t is observable. The model under the alternative is

$$(1-\alpha_{x1}L)(1-\psi_{x1}L)x_t = f_t.$$

Therefore, the null is $H_0 : \psi_{x1} = 0$.[12] Under the alternative, the spectral density of x_t is

$$\frac{\sigma_f^2}{(1-\alpha_{x1}e^{-i\lambda})(1-\alpha_{x1}e^{i\lambda})}\frac{1}{(1-\psi_{x1}e^{-i\lambda})(1-\psi_{x1}e^{i\lambda})}.$$

[12] This is a multiplicative alternative. Instead, we could test $H_0 : \alpha_{x2} = 0$ in the additive alternative

$$(1-\alpha_{x1}L-\alpha_{x2}L^2)x_t = f_t.$$

In that case, it would be more convenient to reparametrise the model in terms of partial autocorrelations as $\alpha_{12} = \alpha_1/(1-\alpha_2)$, $\alpha_{22} = \alpha_2$. We stick to multiplicative alternatives, which cover MA terms too.

The derivative of $G_{xx}(\lambda)$ with respect to ψ_{x1} under the null is

$$\frac{\partial G_{xx}(\lambda)}{\partial \psi_{x1}} = 2(e^{-i\lambda} + e^{i\lambda})\frac{\sigma_f^2}{(1-\alpha_{x1}e^{-i\lambda})(1-\alpha_{x1}e^{i\lambda})} = 2\cos\lambda G_{xx}(\lambda).$$

Hence the spectral version of the score with respect to ψ_{x1} under H_0 is

$$\sum_{j=0}^{T-1} \cos\lambda_j G_{xx}^{-1}(\lambda_j)[2\pi I_{xx}(\lambda_j) - G_{xx}(\lambda_j)] = \sum_{j=0}^{T-1} \cos\lambda_j [2\pi I_{ff}(\lambda_j)],$$

where we have exploited the fact that

$$\sum_{j=0}^{T-1} \frac{\partial G_{xx}(\lambda_j)}{\partial \psi_{x1}} G_{xx}^{-1}(\lambda_j) = \sum_{j=0}^{T-1} \cos\lambda_j = 0.$$

Given that

$$I_{ff}(\lambda_j) = \hat{\gamma}_{ff}(0) + 2\sum_{k=1}^{T-1} \hat{\gamma}_{ff}(k)\cos(k\lambda_j),$$

the spectral version of the score becomes

$$\sum_{j=0}^{T-1} \cos\lambda_j[2\pi I_{ff}(\lambda_j)] = T[\hat{\gamma}_{ff}(1) + \hat{\gamma}_{ff}(T-1)].$$

In turn, the time domain version of the score will be

$$\sum_t (x_t - \alpha_{x1}x_{t-1})(x_{t-1} - \alpha_{x1}x_{t-2}) = \sum_t f_t f_{t-1},$$

which is essentially identical because $\hat{\gamma}_{ff}(T-1) = T^{-1}x_T x_1 = o_p(1)$. Therefore, the LM spectral test is simply checking that the first sample (circulant) autocovariance of f_t coincides with its theoretical value under H_0, exactly like the usual Breusch-Godfrey serial correlation LM test.

3 Neglected serial correlation tests in dynamic factor models

3.1 Testing ARMA(p,q) vs ARMA(p+d,q) (or ARMA(p,q+d)) in the common factor

We can combine our previous results to test the same null hypothesis when x_t is not directly observed. As we saw before, the spectral density of the dynamic GLS estimator of the common factor is

$$G_{x^G x^G}(\lambda) = G_{xx}(\lambda) + [\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda})]^{-1}.$$

As a result,

$$\frac{\partial G_{x^G x^G}(\lambda)}{\partial \psi_{x1}} = \frac{\partial G_{xx}(\lambda)}{\partial \psi_{x1}}.$$

Hence, the score with respect to ψ_{x1} will be given by

$$\frac{1}{2}\sum_{j=0}^{T-1} \frac{\partial G_{xx}(\lambda_j)}{\partial \psi_{x1}} G_{x^G x^G}^{-1}(\lambda_j)[2\pi I_{x^G x^G}(\lambda_j) - G_{x^G x^G}(\lambda_j)].$$

After some straightforward algebraic manipulations, we can show that under the null of $H_0 : \psi_{x1} = 0$ this score can be written as

$$\begin{aligned} & \sum\nolimits_{j=0}^{T-1} \cos \lambda_j G_{xx}^{-1}(\lambda_j)[2\pi I_{x^K x^K}(\lambda_j) - G_{x^K x^K}(\lambda_j)] \\ = & \sum\nolimits_{j=0}^{T-1} \cos \lambda_j [2\pi I_{f^K f^K}(\lambda_j) - G_{f^K f^K}(\lambda_j)]. \end{aligned}$$

Once again, the time domain counterpart to the spectral score with respect to ψ_{x1} is (asymptotically) proportional to the difference between the first sample autocovariance of f_t^K and its theoretical counterpart under H_0. Therefore, the only difference with the observable case is that the autocovariance of f_t^K, which is a forward filter of the Wold innovations of $\mathbf{y}_t$, is no longer 0 when $\psi_{x1} = 0$, although it approaches 0 as the signal to noise ratio increases. In that case, our proposed tests would converge to the usual Breusch-Godfrey LM tests for neglected serial correlation discussed in section 2.7.

Let us illustrate our test by means of a simple example. Imagine that the model under the alternative is:

$$\begin{gathered} \mathbf{y}_t = \boldsymbol{\pi} + \mathbf{c}x_t + \mathbf{u}_t, \quad \mathbf{u}_t = \mathbf{v}_t, \\ (1 - \psi_{x1}L)(1 - \alpha_{x1}L)x_t = f_t, \\ \begin{pmatrix} f_t \\ \mathbf{v}_t \end{pmatrix} |I_{t-1}, \boldsymbol{\theta} \sim N\left[\begin{pmatrix} 0 \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Gamma} \end{pmatrix}\right]. \end{gathered}$$

The results in section 2.6 imply that $x_{t|\infty}^K$ will have the autocorrelation structure of an AR(2) when $\psi_{x1} = 0$, while $f_{t|\infty}^K$ will follow an AR(1) with first order autocovariance $(\mathbf{c}'\mathbf{\Gamma}^{-1}\mathbf{c})\alpha_{x1}/(1 - \alpha_{f^K}^2)$, where

$$\alpha_{f^K} = \frac{1 + \alpha_{x1}^2 + (\mathbf{c}'\mathbf{\Gamma}^{-1}\mathbf{c}) - \sqrt{[(1 + \alpha_{x1})^2 + (\mathbf{c}'\mathbf{\Gamma}^{-1}\mathbf{c})]\,[(1 - \alpha_{x1})^2 + (\mathbf{c}'\mathbf{\Gamma}^{-1}\mathbf{c})]}}{2\alpha_{x1}}.$$

The larger $(\mathbf{c}'\mathbf{\Gamma}^{-1}\mathbf{c})$ is, the closer this autocovariance will be to 0.

The LM test of $H_0 : \psi_{x1} = 0$ will simply compare the first sample autocovariance of $f_{t|\infty}^K$ with its theoretical value above. The advantage of our frequency domain approach is that we obtain those autocovariances without explicitly solving the Riccati equation.

Unfortunately, the approach that we have used to obtain a residual correlation test for the common factor cannot be generally applied to the specific factors since the parameters in $\mathbf{G}_{\mathbf{uu}}(\lambda)$ affect both components of the orthogonalised spectral log-likelihood function. Nevertheless, we can start from first principles by exploiting the fact that

$$\frac{\partial vec[\mathbf{G}_{\mathbf{yy}}(\lambda)]}{\partial \psi_{x1}} = [\mathbf{c}(e^{-i\lambda}) \otimes \mathbf{c}(e^{-i\lambda})]\frac{\partial G_{xx}(\lambda)}{\partial \psi_{x1}}$$

We saw before that under the null of $H_0 : \psi_{x1} = 0$

$$\frac{\partial G_{xx}(\lambda)}{\partial \psi_{x1}} = 2\cos\lambda G_{xx}(\lambda).$$

Not surprisingly, if we introduce these derivatives in the formula for the spectral score with respect to ψ_{x1}, we end up with exactly the same frequency-domain and time-domain expressions.

Empirical researchers often assume that the common factors are white noise for identification purposes, so that $G_{xx}(\lambda) = 1$ under the null. Since we make no assumptions on p and q, our tests trivially apply in that situation too. Similarly, generalisations to test ARMA(p,q) vs ARMA(p+k,q) in the common factor are straightforward, since they only involve higher order autocovariances of $f^K_{t|\infty}$. Similarly, it is easy to show that ARMA(p+k,q) and ARMA(p,q+k) multiplicative alternatives are locally asymptotically equivalent, as in the case of univariate tests for serial correlation in observable time series (see e.g. Godfrey (1988)).[13] Finally, we could also consider (multiplicative) seasonal alternatives.

3.2 Testing ARMA(p,q) vs ARMA(p+d,q) (or ARMA(p,q+d)) in specific factors

Let $\boldsymbol{\psi}'_{\mathbf{u}1} = (\psi_{u_1 1}, \ldots, \psi_{\mathbf{u}_N 1})$. In this case we have that

$$\frac{\partial vec[\mathbf{G_{yy}}(\lambda)]}{\partial \boldsymbol{\psi}'_{\mathbf{u}_1}} = \mathbf{E}_N \frac{\partial vecd[\mathbf{G_{uu}}(\lambda)]}{\partial \boldsymbol{\psi}'_{\mathbf{u}_1}},$$

where $\mathbf{E}_N$ is the "diagonalisation" matrix that maps *vecd* into *vec* (see Magnus (1988)). Straightforward algebraic manipulations allow us to write the score with respect to $\psi_{u_i 1}$ under the null of $H_0 : \boldsymbol{\psi}_{\mathbf{u}_1} = \mathbf{0}$ as

$$\begin{aligned} & \sum\nolimits_{j=0}^{T-1} \cos\lambda_j G^{-1}_{u_i u_i}(\lambda_j)[2\pi I_{u_i^K u_i^K}(\lambda_j) - G_{u_i^K u_i^K}(\lambda_j)] \\ = \; & \sum\nolimits_{j=0}^{T-1} \cos\lambda_j [2\pi I_{v_i^K v_i^K}(\lambda_j) - G_{v_i^K v_i^K}(\lambda_j)]. \end{aligned}$$

Thus, the time domain counterpart to the spectral score with respect to $\psi_{u_i 1}$ will be proportional to the difference between the first sample autocovariance of v^K_{it} and its theoretical value under H_0, as expected. Joint tests that look at several idiosyncratic terms together, as well as the common factor, can be easily obtained by combining the scores involved. As we shall see in sections 4.2 and 5, the component tests are rather good at identifying the source of the rejection.

3.3 Parameter uncertainty

So far we have implicitly assumed known model parameters under the null. In practice, some of them will have to be estimated. Maximum likelihood estimation of the dynamic factor model parameters can be done either in the time domain using the Kalman filter or in the frequency domain.

[13] It would also be possible to develop tests of ARMA(p,q) against ARMA(p+k,q+k) along the lines of Andrews and Ploberger (1996). We leave those tests, which will also depend on the differences between sample and population autocovariances of $f^K_{t|\infty}$, for future research.

The sampling uncertainty surrounding the sample mean $\boldsymbol{\pi}$ is asymptotically inconsequential because the information matrix is block diagonal. The sampling uncertainty surrounding the other parameters is not necessarily so. In fact, block diagonality of the components of the information matrix corresponding to the parameters that define the alternative hypothesis, ψ, and the parameters that define the null, $\boldsymbol{\theta}$, is only obtained in some special cases. One such cases arises when $\mathbf{c}(e^{-i\lambda}) = \mathbf{c}$ and both common and idiosyncratic factors follow AR(1) processes with a common autoregressive coefficient. An important example are the static factor models considered by Fiorentini and Sentana (2012). In that situation, all final prediction errors are white noise, and one can safely ignore the estimation error in $\boldsymbol{\theta}$.

More generally, the solution is the standard one: replace the inverse of the (ψ, ψ) block of the information matrix by the (ψ, ψ) block of the inverse information matrix in the quadratic form that defines the LM test. For this reason, we provide computationally efficient expressions for the scores required to compute the information matrix in Appendix B. Importantly, the dual nature of our proposed tests implies that they can be applied regardless of whether we have estimated the model using a time domain or frequency domain log-likelihood.

4 Monte Carlo simulation

4.1 Size experiment

To evaluate possible finite sample size distortions, we generate 10,000 samples of length 500, plus 50 for initialization, (roughly 4 decades of monthly data). The exact model that we simulate and estimate under the null is

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \end{bmatrix} = \begin{bmatrix} .1 \\ .1 \\ .1 \end{bmatrix} + \begin{bmatrix} 0.7 \\ 0.5 \\ 0.4 \end{bmatrix} x_t + \begin{bmatrix} u_{1,t} \\ u_{2,t} \\ u_{3,t} \end{bmatrix},$$

$$(1 - .4L - .2L^2)x_t = f_t,$$

$$(1 + .4L)u_{1,t} = v_{1,t}, \quad (1 - .6L)u_{2,t} = v_{2,t}, \quad (1 - .2L)u_{3,t} = v_{3,t},$$

$$V(f_t) = 1, \qquad vecd'[V(\mathbf{v}_t)] = (0.4, 0.3, 0.8).$$

We compute LM tests against:

1. First order residual autocorrelation in the common factor (χ^2_1)

2. First and second order residual autocorrelation in the common factor (χ^2_2)

3. First order residual autocorrelation in all the specific factors (χ^2_3)

4. First order residual autocorrelation in common and specific factors (χ^2_4)

Importantly, all our tests are numerically invariant to whether in estimating the model we normalise the variance of the common factor x_t or its innovation f_t to 1 because of the way we compute the information matrix (see Dufour and Dagenais (1991)).

The p-value discrepancy plots (see Davidson and McKinnon (1998)) for the four test are displayed in Figure 1. As can be seen, all tests have virtually no size distortions, with the joint tests showing even smaller distortions than the tests that focus on the common factor only.

4.2 Power experiments

We first simulate and estimate 2,000 samples of length 500, plus 50 for initialization, in which the DGP for the common factors has $\psi_x(L) = (1 - .5L - .25L^2 - .125L^3 - \ldots) = (1 - .5L)^{-1}$ but the same first and second-order autocorrelation as under the null, so that

$$x_t = 0.874x_{t-1} + 0.037x_{t-2} + f_t - 0.5f_{t-1}. \tag{5}$$

We also re-scale the loadings so as to maintain the same unconditional signal to noise ratio as under the null in an attempt to isolate our power results from changes in the degree of observability of the factors. Anything else is left unchanged. The results are reported in Figure 2. As expected, the test that focuses on the correct alternative hypothesis has the largest power, followed by the test that also focuses on second order residual correlation in the common factor, which wastes one degree of freedom. Not surprisingly, the least powerful test is only looking at the specific factors, which nevertheless retains some small power because their estimators are affected by the neglected serial correlation in the common factor.

We also simulate and estimate 2,000 samples of the same length as above in which the DGP for the specific factors has $\psi_{u_i}(L) = (1 + .2L)$, for $i = 1, 2, 3$, but the same first-order autocorrelation as under the null, so that

$$\left.\begin{array}{rcl} u_{1,t} &=& -0.418u_{1,t-1} - 0.044u_{1,t-2} + v_{1,t}, \\ u_{2,t} &=& 0.514u_{2,t-1} + 0.143u_{2,t-2} + v_{2,t}, \\ u_{3,t} &=& 0.185u_{3,t-1} + 0.077u_{3,t-2} + v_{3,t}. \end{array}\right\} \tag{6}$$

Again, we re-scale $V(\mathbf{v}_t)$ in order to match the same unconditional signal to noise ratio as under the null, leaving everything else unchanged. The results are reported in Figure 3. Not surprisingly, the test that focuses on the correct alternative hypothesis has the largest power, followed by the joint test. This time, the tests that focus on the common factor have power essentially equal to nominal size. However, in the experiment reported in Figure 4 we find that when the neglected serial correlation in the specific factor is larger ($\psi_{u_i1} = -.6$) the tests that focus on the common factor regain some non-trivial power because their estimators under the null are contaminated by some of the neglected serial correlation in the specific factors. Nevertheless, given that the rejection rates of the tests that look at the specific terms are essentially 100% for all confidence levels, the results confirm once again that our test correctly identify the source of the rejection.

Finally, we combine the DGPs (5) and (6) above, so that both common and the specific factors are simulated under the alternative. As can be seen in Figure 5 the joint test is the most powerful followed by the test on the specific factors. As expected, though, all test show non-neglegible power in this case.

5 Empirical illustration

We initially replicate the results in Camacho, Pérez-Quirós and Poncela (2012), who construct a monthly US coincident index by combining the indicators of economic activity previously analysed by Stock and Watson (1991), Chauvet (1998) and Chauvet and Pigier (2008). Specifically, they use the industrial production index (IPI), nonfarm payroll employment (EMP), personal income less transfer payments (INC) and real manufacturing and trade sales (SAL). The sample covers the period January 1967 to November 2010 for a total effective sample length of 526 observations. As usual, the seasonally adjusted series are log-transformed and differenced to achieve stationarity. Their basic single factor model specification is

$$\begin{bmatrix} IPI_t \\ EMP_t \\ INC_t \\ SAL_t \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} x_t + \begin{bmatrix} u_{1,t} \\ u_{2,t} \\ u_{3,t} \\ u_{4,t} \end{bmatrix},$$

$$\begin{aligned} x_t &= \phi_{x,1}x_{t-1} + \phi_{x,2}x_{t-2} + f_t, \\ u_{i,t} &= \phi_{i,1}u_{i,t-1} + \phi_{i,2}u_{i,t-2} + v_{i,t}, \qquad i = 1, \dots, 4. \end{aligned}$$

Each variable is individually standardised, the first two observations are discarded and the scale indeterminacy is eliminated by setting $Var(f_t) = 1$. We report the spectral maximum likelihood estimates of the parameters in Table 1.

Table 1: Spectral maximum likelihood estimates

	x	IPI	EMP	INC	SAL
b_i	-	0.68	0.50	0.28	0.45
ϕ_1	0.43	-0.25	0.24	-0.20	-0.36
ϕ_2	0.22	-0.21	0.52	-0.05	-0.16
σ^2	1	0.27	0.25	0.85	0.59

These estimates are very close to the estimates obtained on the basis of the usual time domain log-likelihood.

Our spectral LM test against first order neglected residual serial correlation in the common factor takes the value of 4.28 with a p-value of 3.9%. The same specification test for all three idiosyncratic factors is equal to 34.01, whose p-value is essentially zero. Not surprising, the joint test (36.3) rejects the null of correct specification at all conventional levels of significance.

As we saw in section 4.2, though, the massive rejection of the null in the case of the idiosyncratic factors might partly explain the mild rejection observed for the common factor. For that reason, we re-estimate the model with the same Ar(2) specification for the common factor, but allowing for Arma(2,1) idiosyncratic terms. This time we no longer reject when we look at either the common factor or the idiosyncratic terms.

Camacho, Pérez-Quirós and Poncela (2012) argue that many features of the business cycle are better represented by a Markov switching model than by a linear model. In this regard, we prove in Appendix C that a simple two-regime Markov model for the mean of the common factor would generate the autocorrelation structure of an Arma(1,1) process. Therefore, the Ar(2) specification for x_t should have been rejected. Nevertheless, it is conceptually possible that the implied Arma(1,1) process could be such that an Ar(2) still provides a reasonable linear approximation. In any case, our results suggest that their Markov switching model should allow for more flexible dynamics in the idiosyncratic terms.

6 Conclusions and extensions

We derive computationally simple and intuitive expressions for score tests of neglected serial correlation in common and idiosyncratic factors in dynamic factor models using frequency domain methods. Our tests can focus on all state variables, the common factors only, the specific factors only, or indeed some of their elements. The implicit orthogonality conditions are analogous to the conditions obtained by treating the Wiener-Kolmogorov-Kalman smoothed estimators of the innovations in common and idiosyncratic factors as if they were observed, but they account for their final estimation errors.

We conduct Monte Carlo exercises to study the finite sample reliability and power of our proposed tests. Our simulation results suggest that they have rather accurate sizes in finite samples. They also confirm that our tests have power to detect neglected serial correlation in common or specific factors, and that they are also systematically able to correctly identify the source of the rejection.

Finally, we evaluate the empirical usefulness of our tests by assessing the dynamic factor model used by Camacho, Pérez-Quirós and Poncela (2012) to construct a coincident indicator for the US.

The testing procedures developed in the previous sections can be extended in several interesting directions. One obvious possibility would be to consider models with multiple common factors. Although this would be intensive in notation, such an extension would be otherwise straightforward after dealing with the usual identification issues before estimating the model under the null. It should also be possible to extend our procedures to multivariate regressions whose residuals follow a dynamic factor model. In that regard, Fiorentini and Sentana (2012) provide a thorough comparison of the LM tests that we have considered with serial correlation tests based on reduced form residuals, showing that there are clear power gains from exploiting the cross-sectional dependence structure implicit in factor models.

Another worthwhile extension would cover restrictions on the dynamic factor loadings. Examples of interesting null hypotheses of this sort would be the equality of the impulse responses of a common factor for two or more observed series, or a finite lag limit for those responses. Given that we show in Appendix B that the scores of the dynamic loadings can be related to the normal equations in a distributed lag regression of $\mathbf{y}_t$ on $x^K_{t|\infty}$, it should be fairly straightforward to derive those tests. Relatedly, we could also study the asymptotic power properties of our proposed tests against such alternatives, or indeed alternatives in which there are missing dynamic factors under the null.

Throughout the paper we have maintained the assumption of normality. To understand its implications, let $\boldsymbol{\mu}_t$ and $\boldsymbol{\Sigma}_t$ denote the conditional mean vector and covariance matrix of $\mathbf{y}_t$ given its past alone, which can be obtained from the prediction equations of the Kalman filter. Given that the serial correlation parameters ψ effectively enter through $\boldsymbol{\mu}_t$ only, the information matrix equality should continue to hold for their scores. In any case, Dunsmuir and Hannan (1976) and Dunsmuir (1979) already provided sandwich formulas for the asymptotic variances of estimators obtained by maximising the spectral log-likelihood function (3). Similarly, it would be straightforward to exploit the asymptotic orthogonality of the frequency components of the Whittle likelihood to devise suitable bootstrap procedures (see Dahlhaus and Janas (1996) or Kirch and Politis (2011)).

Although we have only considered state variables with rational spectral densities, in principle our methods could be applied to long memory processes too. In this regard, it would be worth exploring the long memory alternatives considered by Robinson (1991). More generally, it would also be interesting to consider non-parametric alternatives such as the ones studied by Hong (1996), in which the lag length is implicitly determined by the choice of bandwidth parameter in a kernel-based estimator of a spectral density matrix. Another potential extension would directly deal with non-stationary factor models, such as the common stochastic trends models in Peña and Poncela (2006), without transforming the observed variables to achieve stationarity. In this regard, we would expect our proposed tests to remain valid in those circumstances too because they focus on the stationary components of dynamic factors models.

Given their ubiquitousness in the recent empirical literature (see e.g. Bai and Ng (2008) and the references therein), the extension of our methods to approximate factor models in which (i) the cross-sectional dimension is non-negligible relative to the time series dimension; and (ii) there is some mild contemporaneous and dynamic correlation between the idiosyncratic terms would constitute a very valuable addition. Although Doz, Giannone and Reichlin (2012) have recently proved the consistency of the Gaussian pseudo ML estimators that we have used in such contexts, the extension of our tests would require asymptotic distributions under suitable rates, which are as yet unknown.

Finally, it is worth mentioning that although we have exploited some specificities of dynamic factor models, our procedures can be easily extended to most unobserved components time series processes, and in particular to UCARIMA models (see Maravall (1999)) and the state-space models underlying the recent nowcasting literature (see Banbura, Giannone and Reichlin (2010) and the references therein). We are currently pursuing some of these research avenues.

References

Andrews, D.W.K. and Ploberger, W. (1996): "Testing for serial correlation against an ARMA(1,1) process", *Journal of the American Statistical Association* 91, 1331–1342.

Bai, J. and Ng, S. (2008): "Large dimensional factor analysis", *Foundations and Trends in Econometrics* 3, 89–163.

Banbura, M., Giannone, D., and Reichlin, L. (2010): "Nowcasting", forthcoming in M.P. Clemens and D.F. Hendry (eds.), *Oxford Handbook on Economic Forecasting.*

Breusch, T. S. and Pagan, A.R. (1980): "The Lagrange Multiplier test and its applications to model specification in econometrics", *Review of Economic Studies* 47, 239-253.

Camacho, M., Pérez-Quirós, G. and Poncela, P. (2012), "Extracting nonlinear signals from several economic indicators", *CEPR Discussion Paper* 8865.

Chauvet, M. (1998): "An econometric characterization of business cycle dynamics with factor structure and regime switches", *International Economic Review* 39, 969-96.

Chauvet, M., and Piger, J. (2008): "A comparison of the real-time performance of business cycle dating methods", *Journal of Business and Economic Statistics* 26, 42-49.

Choudhuri , N., Ghosal S. and Roy, A. (2004): "Contiguity of the Whittle measure for a Gaussian time series", *Biometrika* 91, 211-218.

Connor, G., Goldberg, L.R. Korajczik, R.A. (2010): *Portfolio risk analysis*, Princeton.

Dahlhaus, R. and Janas, D. (1996): "A frequency domain bootstrap for ratio statistics in time series analysis", *Annals of Statistics* 24, 1934-1963.

Davidson, R. and MacKinnon, J.G. (1998): "Graphical methods for investigating the size and power of tests statistics", *The Manchester School* 66, 1-26.

Dempster, A., Laird, N. and Rubin, D. (1977): "Maximum likelihood from incomplete data via the EM algorithm", *Journal of the Royal Statistical Society B* 39, 1-38.

Diebold, F.X. and Rudebusch, G.D. (1996), "Measuring business cycles: a modern perspective", *Review of Economics and Statistics*, 78, 67-77.

Diebold, F.X., Rudebusch, G.D. and Aruoba, B. (2006), "The macroeconomy and the yield curve: a dynamic latent factor approach", *Journal of Econometrics* 131, 309-338.

Doz, C., Giannone, D. and Reichlin, L. (2012): "A quasi maximum likelihood approach for large approximate dynamic factor models", *Review of Economics and Statistics* 94, 1014-1024.

Dufour, J.M. and Dagenais, M.G. (1991): "Invariance, nonlinear models and asymptotic tests", *Econometrica* 59, 1601-1615.

Dungey, M., Martin, V.L. and Pagan, A.R. (2000): "A multivariate latent factor decomposition of international bond yield spreads", *Journal of Applied Econometrics* 15, 697-715.

Dunsmuir, W. (1979): "A central limit theorem for parameter estimation in stationary vector time series and its application to models for a signal observed with noise", *Annals of Statistics* 7, 490-506.

Dunsmuir, W. and Hannan, E.J. (1976): "Vector linear time series models", *Advances in Applied Probability* 8, 339-364.

Engle, R.F. and Watson, M.W. (1980): "Formulation générale et estimation de models multidimensionnels temporels a facteurs explicatifs non observables", *Cahiers du Séminaire d'Économétrie* 22, 109-125.

Engle, R.F. and Watson, M.W. (1981): "A one-factor multivariate time series model of metropolitan wage rates", *Journal of the American Statistical Association* 76, 774-781.

Fernández, F.J. (1990): "Estimation and testing of a multivariate exponential smoothing model", *Journal of Time Series Analysis* 11, 89–105.

Fiorentini, G. (1995): *Conditional heteroskedasticity: some results on estimation, inference and signal extraction, with an application to seasonal adjustment*, unpublished Doctoral Dissertation, European University Institute.

Fiorentini, G. and Sentana, E. (2009): "Dynamic specification tests for static factor models", CEMFI Working Paper 0912.

Fiorentini, G. and Sentana, E. (2012): "Tests for serial dependence in static, non-Gaussian factor models", CEMFI Working Paper 1211.

Fiorentini, G., Sentana, E. and Shephard, N. (2004): "Likelihood estimation of latent generalised ARCH structures", *Econometrica* 72, 1481-1517.

Geweke, J.F. (1977): "The dynamic factor analysis of economic time series models," in D. Aigner and A. Goldberger (eds.), *Latent Variables in Socioeconomic Models*, 365-383, North-Holland.

Geweke, J.F. and Singleton, K.J (1981): "Maximum likelihood "confirmatory" factor analysis of economic time series", *International Economic Review* 22, 37-54.

Godfrey, L.G. (1988): *Misspecification tests in econometrics: the Lagrange multiplier principle and other approaches*, Econometric Society Monographs.

Gómez, V. (1999): "Three equivalent methods for filtering finite nonstationary time series", *Journal of Business and Economic Statistics* 17, 109-116.

Gourieroux, C., Monfort, A. and Renault, E. (1991): "A general framework for factor models", mimeo, INSEE.

Gregory, A.W., Head, A.C. and Raynauld, J. (1997): "Measuring world business cycles", *International Economic Review* 38, 677-701.

Harvey, A.C. (1989): *Forecasting, structural models and the Kalman filter*, Cambridge University Press, Cambridge.

Heaton, C. and Solo, V. (2004): "Identification of causal factor models of stationary time series", *Econometrics Journal* 7, 618-627.

Hong, Y. (1996): "Consistent testing for serial correlation of unknown form" *Econometrica* 64, 837-864.

Jegadeesh, N. and G.G. Pennacchi (1996): "The behavior of interest rates implied by the term structure of eurodollar futures", *Journal of Money, Credit and Banking* 28, 426-446.

Jungbacker, B. and S.J. Koopman (2008): "Likelihood-based analysis for dynamic factor models", Tinbergen Institute Discussion Paper 2008-0007.

Kirch, C. and Politis, D.N. (2011): "TFT-Bootstrap: Resampling time series in the frequency domain to obtain replicates in the time domain", *Annals of Statistics* 39, 1427-1470.

Landefeld, J.S., Seskin, E.P. and Fraumeni, B.M. (2008): "Taking the pulse of the economy: measuring GDP", *Journal of Economic Perspectives* 22, 193–216.

Lawley, D.N. and Maxwell, A.E. (1971): *Factor analysis as a statistical method*, 2nd ed., Butterworths, London.

Litterman, R. and Sargent, T.J. (1979): "Detecting neutral price level changes and the effects of aggregate demand with index models", Research Department Working Paper 125, Federal Reserve Bank of Minneapolis.

Magnus, J.R. (1988): *Linear structures*, Oxford University Press, New York.

Magnus, J.R. and Neudecker, H. (1988): *Matrix differential calculus with applications in Statistics and Econometrics*, Wiley, Chichester

Maravall, A. (1999): "Unobserved components in economic time series" in M.H. Pesaran and M.R. Wickens (eds.), *Handbook of Applied Econometrics. Volume I: Macroeconomics*, Blackwell.

Mody, A. and Taylor, M.P. (2007): "Regional vulnerability: The case of East Asia", *Journal of International Money and Finance* 26, 1292-1310.

Newey, W.K. (1985): "Maximum likelihood specification testing and conditional moment tests", *Econometrica* 53, 1047-70.

Peña, D. and Box, G.E.P. (1987): "Identifying a simplifying structure in time series", *Journal of the American Statistical Association* 82, 836-843.

Peña, D. and Poncela, P. (2006): "Nonstationary dynamic factor analysis", *Journal of Statistical Planning and Inference* 136, 1237-1257.

Quah, D. and Sargent, T. (1993): "A dynamic index model for large cross sections", in J.H. Stock and M.W. Watson (eds.), *Business cycles, indicators and forecasting*, 285-310, University of Chicago Press.

Robinson, P. M. (1991): "Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression", *Journal of Econometrics* 47, 67-84.

Sentana, E. (2000): "The likelihood function of conditionally heteroskedastic factor models", *Annales d'Economie et de Statistique* 58, 1-19.

Sentana, E. (2004): "Factor representing portfolios in large asset markets", *Journal of Econometrics* 119, 257-289.

Singleton, K.J. (1981): "A latent time series model of the cyclical behavior of interest rates", *International Economic Review* 21, 559-575.

Stock, J.H. and Watson, M.W. (1989): "New indexes of coincident and leading economic indicators", *NBER Macroeconomics Annual* 4, 351-394.

Stock, J., and Watson, M.W. (1991): "A probability model of the coincident economic indicators", in K. Lahiri and G. Moore (eds.), *Leading economic indicators: new approaches and forecasting records*, Cambridge University Press.

Stock, J.H. and Watson, M.W. (1993): "A procedure for predicting recessions with leading indicators: econometric issues and recent experience", in J.H. Stock and M.W. Watson (eds.), *Business cycles, indicators, and forecasting*, 95-153, University of Chicago Press.

Tauchen, G. (1985): "Diagnostic testing and evaluation of maximum likelihood models", *Journal of Econometrics* 30, 415-443.

Watson, M.W. and Engle, R.F. (1983): "Alternative algorithms for estimation of dynamic MIMIC, factor, and time varying coefficient regression models", *Journal of Econometrics* 23, 385-400.

Watson, M.W. and Kraft, D.F. (1984): "Testing the interpretation of indices in a macroeconomic index model", *Journal of Monetary Economics* 13, 165-182.

Whittle, P. (1962): "Gaussian estimation in stationary time series", *Bulletin of the International Statistical Institute* 39, 105-129.

Appendix

A Time domain tests

The simplest state space representation of a dynamic factor model with an AR(2) common factor is:

1. Measurement equation:
$$\mathbf{y}_t = (\mathbf{c}|\mathbf{0}) \begin{pmatrix} x_t \\ x_{t-1} \end{pmatrix} + \mathbf{u}_t,$$

2. Transition equation:
$$\begin{pmatrix} x_t \\ x_{t-1} \end{pmatrix} = \begin{pmatrix} \rho_1 & \rho_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ x_{t-2} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} f_t.$$

Therefore, the Kalman filter prediction equations will be:

$$\begin{pmatrix} x_{t|t-1} \\ x_{t-1|t-1} \end{pmatrix} = \begin{pmatrix} \rho_1 & \rho_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_{t-1|t-1} \\ x_{t-2|t-1} \end{pmatrix} = \begin{pmatrix} \rho_1 x_{t-1|t-1} + \rho_2 x_{t-2|t-1} \\ x_{t-1|t-1} \end{pmatrix}, \tag{A1}$$

$$\boldsymbol{\mu}_t(\boldsymbol{\theta}) = \mathbf{y}_{t|t-1} = (\mathbf{c}|\mathbf{0}) \begin{pmatrix} x_{t|t-1} \\ x_{t-1|t-1} \end{pmatrix} = \mathbf{c} x_{t|t-1},$$

$$\begin{aligned} \boldsymbol{\Omega}_{t|t-1} &= \begin{pmatrix} \omega_{11t|t-1} & \omega_{21t|t-1} \\ \omega_{21t|t-1} & \omega_{22t|t-1} \end{pmatrix} \\ &= \begin{pmatrix} \rho_1 & \rho_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \omega_{11t-1|t-1} & \omega_{21t-1|t-1} \\ \omega_{21t-1|t-1} & \omega_{22t-1|t-1} \end{pmatrix} \begin{pmatrix} \rho_1 & 1 \\ \rho_2 & 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} 1 \begin{pmatrix} 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} \rho_1^2 \omega_{11t-1|t-1} + 2\rho_1\rho_2\omega_{21t-1|t-1} + \rho_2^2\omega_{22t-1|t-1} + 1 & \rho_1\omega_{11t-1|t-1} + \rho_2\omega_{21t-1|t-1} \\ \rho_1\omega_{11t-1|t-1} + \rho_2\omega_{21t-1|t-1} & \omega_{11t-1|t-1} \end{pmatrix} \end{aligned} \tag{A2}$$

and

$$\boldsymbol{\Sigma}_{t|t-1}(\boldsymbol{\theta}) = (\mathbf{c}|\mathbf{0})\, \boldsymbol{\Omega}_{t|t-1} \begin{pmatrix} \mathbf{c}' \\ \mathbf{0}' \end{pmatrix} + \boldsymbol{\Gamma} = \mathbf{c}\omega_{11t|t-1}\mathbf{c}' + \boldsymbol{\Gamma}.$$

Similarly, the updating equations will be:

$$\begin{aligned} \begin{pmatrix} x_{t|t} \\ x_{t-1|t} \end{pmatrix} &= \begin{pmatrix} x_{t|t-1} \\ x_{t-1|t-1} \end{pmatrix} + \begin{pmatrix} \omega_{11t|t-1} & \omega_{21t|t-1} \\ \omega_{21t|t-1} & \omega_{22t|t-1} \end{pmatrix} \begin{pmatrix} \mathbf{c}' \\ \mathbf{0}' \end{pmatrix} \boldsymbol{\Sigma}^{-1}_{t|t-1}(\boldsymbol{\theta})(\mathbf{y}_t - \mathbf{c}x_{t|t-1}) \\ &= \begin{pmatrix} x_{t|t-1} + \omega_{11t|t-1}\mathbf{c}'\boldsymbol{\Sigma}^{-1}_{t|t-1}(\boldsymbol{\theta})(\mathbf{y}_t - \mathbf{c}x_{t|t-1}) \\ x_{t-1|t-1} + \omega_{21t|t-1}\mathbf{c}'\boldsymbol{\Sigma}^{-1}_{t|t-1}(\boldsymbol{\theta})(\mathbf{y}_t - \mathbf{c}x_{t|t-1}) \end{pmatrix} \end{aligned} \tag{A3}$$

and

$$\begin{aligned} \boldsymbol{\Omega}_{t|t} &= \boldsymbol{\Omega}_{t|t-1} - \begin{pmatrix} \omega_{11t|t-1} & \omega_{21t|t-1} \\ \omega_{21t|t-1} & \omega_{22t|t-1} \end{pmatrix} \begin{pmatrix} \mathbf{c}' \\ \mathbf{0}' \end{pmatrix} \boldsymbol{\Sigma}^{-1}_{t|t-1}(\boldsymbol{\theta})\,(\mathbf{c}|\mathbf{0}) \begin{pmatrix} \omega_{11t|t-1} & \omega_{21t|t-1} \\ \omega_{21t|t-1} & \omega_{22t|t-1} \end{pmatrix} \\ &= \begin{pmatrix} \omega_{11t|t-1} - \omega^2_{11t|t-1}\mathbf{c}'\boldsymbol{\Sigma}^{-1}_{t|t-1}(\boldsymbol{\theta})\mathbf{c} & \omega_{21t|t-1} - \omega_{21t|t-1}\omega_{11t|t-1}\mathbf{c}'\boldsymbol{\Sigma}^{-1}_{t|t-1}(\boldsymbol{\theta})\mathbf{c} \\ \omega_{21t|t-1} - \omega_{21t|t-1}\omega_{11t|t-1}\mathbf{c}'\boldsymbol{\Sigma}^{-1}_{t|t-1}(\boldsymbol{\theta})\mathbf{c} & \omega_{22t|t-1} - \omega^2_{21t|t-1}\mathbf{c}'\boldsymbol{\Sigma}^{-1}_{t|t-1}(\boldsymbol{\theta})\mathbf{c} \end{pmatrix}. \end{aligned} \tag{A4}$$

If we call

$$f_{t|t} = \mathbf{c}'\mathbf{\Sigma}^{-1}_{t|t-1}(\boldsymbol{\theta})(\mathbf{y}_t - \mathbf{c}x_{t|t-1}) \tag{A5}$$

and

$$\varpi_{t|t} = 1 - \omega_{11t|t-1}\mathbf{c}'\mathbf{\Sigma}^{-1}_{t|t-1}(\boldsymbol{\theta})\mathbf{c}, \tag{A6}$$

where we can interpret $f_{t|t}$ as the conditional expectation of f_t given $y_t, y_{t-1}, \ldots$, and $\varpi_{t|t}$ as the covariance between x_t and f_t conditional on $y_t, y_{t-1}, \ldots$, then we can write the previous expressions as

$$\begin{pmatrix} x_{t|t} \\ x_{t-1|t} \end{pmatrix} = \begin{pmatrix} x_{t|t-1} + \omega_{11t|t-1}f_{t|t} \\ x_{t-1|t-1} + \omega_{21t|t-1}f_{t|t} \end{pmatrix}$$

and

$$\mathbf{\Omega}_{t|t} = \begin{pmatrix} \omega_{11t|t-1}\varpi_{t|t} & \omega_{21t|t-1}\varpi_{t|t} \\ \omega_{21t|t-1}\varpi_{t|t} & (\omega_{22t|t-1} - \omega^2_{21t|t-1}\omega^{-1}_{11t|t-1}) + \omega^2_{21t|t-1}\omega^{-1}_{11t|t-1}\varpi_{t|t} \end{pmatrix}.$$

These expressions simplify considerably further when $|\mathbf{\Gamma}| > 0$, in which case the Woodbury formula yields

$$\mathbf{\Sigma}^{-1}_{t|t-1}(\boldsymbol{\theta}) = \mathbf{\Gamma}^{-1} - \frac{\mathbf{\Gamma}^{-1}\mathbf{c}\mathbf{c}'\mathbf{\Gamma}^{-1}}{\omega^{-1}_{11t|t-1} + \mathbf{c}'\mathbf{\Gamma}^{-1}\mathbf{c}}.$$

Specifically,

$$\varpi_{t|t} = \frac{\omega^{-1}_{11t|t-1}}{\omega^{-1}_{11t|t-1} + \mathbf{c}'\mathbf{\Gamma}^{-1}\mathbf{c}}$$

and

$$f_{t|t} = \varpi_{t|t}\mathbf{c}'\mathbf{\Gamma}^{-1}(\mathbf{y}_t - \mathbf{c}x_{t|t-1}),$$

where we have used the fact that

$$\mathbf{c}'\mathbf{\Sigma}^{-1}_{t|t-1}(\boldsymbol{\theta}) = \mathbf{c}'\mathbf{\Gamma}^{-1} - \frac{\mathbf{c}'\mathbf{\Gamma}^{-1}\mathbf{c}\mathbf{c}'\mathbf{\Gamma}^{-1}}{\omega^{-1}_{11t|t-1} + \mathbf{c}'\mathbf{\Gamma}^{-1}\mathbf{c}} = \frac{\omega^{-1}_{11t|t-1}\mathbf{c}'\mathbf{\Gamma}^{-1}}{\omega^{-1}_{11t|t-1} + \mathbf{c}'\mathbf{\Gamma}^{-1}\mathbf{c}}$$

and

$$\mathbf{c}'\mathbf{\Sigma}^{-1}_{t|t-1}(\boldsymbol{\theta})\mathbf{c} = \frac{\omega^{-1}_{11t|t-1}\mathbf{c}'\mathbf{\Gamma}^{-1}\mathbf{c}}{\omega^{-1}_{11t|t-1} + \mathbf{c}'\mathbf{\Gamma}^{-1}\mathbf{c}}.$$

In order to find the log-likelihood score, it is convenient to write

$$\begin{aligned} d\boldsymbol{\mu}_t(\boldsymbol{\theta}) &= d\mathbf{c} \cdot x_{t|t-1} + \mathbf{c} \cdot dx_{t|t-1}, \\ d\mathbf{\Sigma}_{t|t-1}(\boldsymbol{\theta}) &= d\mathbf{c} \cdot \omega_{11t|t-1}\mathbf{c}' + \mathbf{c} \cdot d\omega_{11t|t-1} \cdot \mathbf{c}' + \mathbf{c}\omega_{11t|t-1} \cdot d\mathbf{c}' + d\mathbf{\Gamma}, \end{aligned}$$

whence

$$\begin{aligned} \frac{\partial \boldsymbol{\mu}_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} &= (x_{t|t-1} \otimes \mathbf{I}_N)\frac{\partial \mathbf{c}}{\partial \boldsymbol{\theta}'} + \mathbf{c}\frac{\partial x_{t|t-1}}{\partial \boldsymbol{\theta}'}, \\ \frac{\partial vec[\mathbf{\Sigma}_t(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}'} &= (\mathbf{I}_{N^2} \otimes \mathbf{K}_{NN})(\mathbf{c}\omega_{11t|t-1} \otimes \mathbf{I}_N)\frac{\partial \mathbf{c}}{\partial \boldsymbol{\theta}'} + (\mathbf{c} \otimes \mathbf{c})\frac{\partial \omega_{11t|t-1}}{\partial \boldsymbol{\theta}'} + \mathbf{E}_N\frac{\partial \gamma}{\partial \boldsymbol{\theta}'}, \\ \frac{\partial \begin{pmatrix} x_{t|t-1} \\ x_{t-1|t-1} \end{pmatrix}}{\partial \boldsymbol{\theta}'} &= [\begin{pmatrix} x_{t-1|t-1} & x_{t-2|t-1} \end{pmatrix} \otimes \mathbf{I}_2]\frac{\partial vec\begin{pmatrix} \rho_1 & \rho_2 \\ 1 & 0 \end{pmatrix}}{\partial \boldsymbol{\theta}'} + \begin{pmatrix} \rho_1 & \rho_2 \\ 1 & 0 \end{pmatrix}\frac{\partial \begin{pmatrix} x_{t-1|t-1} \\ x_{t-2|t-1} \end{pmatrix}}{\partial \boldsymbol{\theta}'} \end{aligned}$$

and

$$\frac{\partial vec(\mathbf{\Omega}_{t|t-1})}{\partial \boldsymbol{\theta}'} = (\mathbf{I}_4 \otimes \mathbf{K}_{22}) \left[\begin{pmatrix} \rho_1 & \rho_2 \\ 1 & 0 \end{pmatrix} \mathbf{\Omega}_{t-1|t-1} \otimes \mathbf{I}_2 \right] \frac{\partial vec \begin{pmatrix} \rho_1 & \rho_2 \\ 1 & 0 \end{pmatrix}}{\partial \boldsymbol{\theta}'} + \left[\begin{pmatrix} \rho_1 & \rho_2 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} \rho_1 & \rho_2 \\ 1 & 0 \end{pmatrix} \right] \frac{\partial vec(\mathbf{\Omega}_{t-1|t-1})}{\partial \boldsymbol{\theta}'}.$$

These last two expressions can be considerably simplified if we differentiate (A1) and (A2) directly. Specifically,

$$\frac{\partial x_{t|t-1}}{\partial \boldsymbol{\theta}'} = x_{t-1|t-1}\frac{\partial \rho_1}{\partial \boldsymbol{\theta}'} + x_{t-2|t-1}\frac{\partial \rho_2}{\partial \boldsymbol{\theta}'} + \rho_1 \frac{\partial x_{t-1|t-1}}{\partial \boldsymbol{\theta}'} + \rho_2 \frac{\partial x_{t-2|t-1}}{\partial \boldsymbol{\theta}'},$$

$$\begin{aligned} \frac{\partial \omega_{11t|t-1}}{\partial \boldsymbol{\theta}'} &= 2(\rho_1 \omega_{11t-1|t-1} + \rho_2 \omega_{21t-1|t-1})\frac{\partial \rho_1}{\partial \boldsymbol{\theta}'} + 2(\rho_1 \omega_{21t-1|t-1} + \rho_2 \omega_{22t-1|t-1})\frac{\partial \rho_2}{\partial \boldsymbol{\theta}'} \\ &\quad + \rho_1^2 \frac{\partial \omega_{11t-1|t-1}}{\partial \boldsymbol{\theta}'} + 2\rho_1\rho_2 \frac{\partial \omega_{21t-1|t-1}}{\partial \boldsymbol{\theta}'} + \rho_2^2 \frac{\partial \omega_{22t-1|t-1}}{\partial \boldsymbol{\theta}'}, \\ \frac{\partial \omega_{21t|t-1}}{\partial \boldsymbol{\theta}'} &= \omega_{11t-1|t-1}\frac{\partial \rho_1}{\partial \boldsymbol{\theta}'} + \omega_{21t-1|t-1}\frac{\partial \rho_2}{\partial \boldsymbol{\theta}'} + \rho_1 \frac{\partial \omega_{11t-1|t-1}}{\partial \boldsymbol{\theta}'} + \rho_2 \frac{\partial \omega_{21t-1|t-1}}{\partial \boldsymbol{\theta}'}, \end{aligned}$$

and

$$\frac{\partial \omega_{22t|t-1}}{\partial \boldsymbol{\theta}'} = \frac{\partial \omega_{11t-1|t-1}}{\partial \boldsymbol{\theta}'}.$$

In any case, we require expressions for

$$\frac{\partial \begin{pmatrix} x_{t-1|t-1} & x_{t-2|t-1} \end{pmatrix}}{\partial \boldsymbol{\theta}}$$

and

$$\frac{\partial vec'(\mathbf{\Omega}_{t-1|t-1})}{\partial \boldsymbol{\theta}},$$

which we can obtain by differentiating (A3) and (A4). Specifically,

$$\begin{aligned} dx_{t|t} &= dx_{t|t-1} + d\omega_{11t|t-1} \cdot \mathbf{c}'\mathbf{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})(\mathbf{y}_t - \mathbf{c}x_{t|t-1}) + \omega_{11t|t-1} \cdot d\mathbf{c}' \cdot \mathbf{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})(\mathbf{y}_t - \mathbf{c}x_{t|t-1}) \\ &\quad - \omega_{11t|t-1}\mathbf{c}'\mathbf{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta}) \cdot d\mathbf{\Sigma}_{t|t-1}(\boldsymbol{\theta}) \cdot \mathbf{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})(\mathbf{y}_t - \mathbf{c}x_{t|t-1}) \\ &\quad - \omega_{11t|t-1}\mathbf{c}'\mathbf{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta}) \cdot d\mathbf{c} \cdot x_{t|t-1} - \omega_{11t|t-1}\mathbf{c}'\mathbf{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c} \cdot dx_{t|t-1} \end{aligned}$$

and

$$\begin{aligned} dx_{t-1|t} &= dx_{t-1|t-1} + d\omega_{21t|t-1} \cdot \mathbf{c}'\mathbf{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})(\mathbf{y}_t - \mathbf{c}x_{t|t-1}) + \omega_{21t|t-1} \cdot d\mathbf{c}' \cdot \mathbf{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})(\mathbf{y}_t - \mathbf{c}x_{t|t-1}) \\ &\quad - \omega_{21t|t-1}\mathbf{c}'\mathbf{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta}) \cdot d\mathbf{\Sigma}_{t|t-1}(\boldsymbol{\theta}) \cdot \mathbf{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})(\mathbf{y}_t - \mathbf{c}x_{t|t-1}) \\ &\quad - \omega_{21t|t-1}\mathbf{c}'\mathbf{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta}) \cdot d\mathbf{c} \cdot x_{t|t-1} - \omega_{21t|t-1}\mathbf{c}'\mathbf{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c} \cdot dx_{t|t-1}, \end{aligned}$$

whence

$$\frac{\partial x_{t|t}}{\partial \boldsymbol{\theta}'} = \frac{\partial x_{t|t-1}}{\partial \boldsymbol{\theta}'} + \mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})(\mathbf{y}_t - \mathbf{c}x_{t|t-1})\frac{\partial \omega_{11t|t-1}}{\partial \boldsymbol{\theta}'} + \omega_{11t|t-1}(\mathbf{y}_t - \mathbf{c}x_{t|t-1})'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\frac{\partial \mathbf{c}}{\partial \boldsymbol{\theta}'}$$
$$-[(\mathbf{y}_t - \mathbf{c}x_{t|t-1})'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta}) \otimes \omega_{11t|t-1}\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})]\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}'}$$
$$-x_{t|t-1}\omega_{11t|t-1}\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\frac{\partial \mathbf{c}}{\partial \boldsymbol{\theta}'} - \omega_{11t|t-1}\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c}\frac{\partial x_{t|t-1}}{\partial \boldsymbol{\theta}'} \tag{A7}$$

and

$$\frac{\partial x_{t-1|t}}{\partial \boldsymbol{\theta}'} = \frac{\partial x_{t-1|t-1}}{\partial \boldsymbol{\theta}'} + \mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})(\mathbf{y}_t - \mathbf{c}x_{t|t-1})\frac{\partial \omega_{21t|t-1}}{\partial \boldsymbol{\theta}'} + \omega_{21t|t-1}(\mathbf{y}_t - \mathbf{c}x_{t|t-1})'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\frac{\partial \mathbf{c}}{\partial \boldsymbol{\theta}'}$$
$$-[(\mathbf{y}_t - \mathbf{c}x_{t|t-1})'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta}) \otimes \omega_{21t|t-1}\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})]\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}'}$$
$$-x_{t|t-1}\omega_{21t|t-1}\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\frac{\partial \mathbf{c}}{\partial \boldsymbol{\theta}'} - \omega_{21t|t-1}\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c}\frac{\partial x_{t|t-1}}{\partial \boldsymbol{\theta}'}.$$

Similarly,

$$d\omega_{11t|t} = (1 - 2\omega_{11t|t-1}\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c})d\omega_{11t|t-1} - \omega_{11t|t-1}^2 \cdot d\mathbf{c}' \cdot \boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c}$$
$$+\omega_{11t|t-1}^2\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta}) \cdot d\boldsymbol{\Sigma}_{t|t-1}(\boldsymbol{\theta}) \cdot \boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c} - \omega_{11t|t-1}^2\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta}) \cdot d\mathbf{c},$$
$$d\omega_{21t|t} = (1 - \omega_{11t|t-1}\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c})d\omega_{21t|t-1} - \omega_{21t|t-1} \cdot d\omega_{11t|t-1} \cdot \mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c}$$
$$-\omega_{21t|t-1}\omega_{11t|t-1} \cdot d\mathbf{c}' \cdot \boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c} + \omega_{21t|t-1}\omega_{11t|t-1}\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta}) \cdot d\boldsymbol{\Sigma}_{t|t-1}(\boldsymbol{\theta}) \cdot \boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c}$$
$$-\omega_{21t|t-1}\omega_{11t|t-1}\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta}) \cdot d\mathbf{c}$$

and

$$\begin{aligned} d\omega_{22t|t} &= d\omega_{22t|t-1} - 2\omega_{21t|t-1}\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c}d\omega_{21t|t-1} - \omega_{21t|t-1}^2 \cdot d\mathbf{c}' \cdot \boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c} \\ &\quad +\omega_{21t|t-1}^2\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta}) \cdot d\boldsymbol{\Sigma}_{t|t-1}(\boldsymbol{\theta}) \cdot \boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c} - \omega_{21t|t-1}^2\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta}) \cdot d\mathbf{c}, \end{aligned}$$

whence

$$\frac{\partial \omega_{11t|t}}{\partial \boldsymbol{\theta}'} = (1 - 2\omega_{11t|t-1}\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c})\frac{\partial \omega_{11t|t-1}}{\partial \boldsymbol{\theta}'} - 2\omega_{11t|t-1}^2\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\frac{\partial \mathbf{c}}{\partial \boldsymbol{\theta}'}$$
$$+[\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta}) \otimes \omega_{11t|t-1}^2\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})]\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}'} \tag{A8}$$
$$\frac{\partial \omega_{21t|t}}{\partial \boldsymbol{\theta}'} = (1 - \omega_{11t|t-1}\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c})\frac{\partial \omega_{21t|t-1}}{\partial \boldsymbol{\theta}'} - \omega_{21t|t-1}\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c}\frac{\partial \omega_{11t|t-1}}{\partial \boldsymbol{\theta}'}$$
$$-2\omega_{21t|t-1}\omega_{11t|t-1}\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\frac{\partial \mathbf{c}}{\partial \boldsymbol{\theta}'} + [\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta}) \otimes \omega_{21t|t-1}\omega_{11t|t-1}\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})]\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}'}$$

and

$$\begin{aligned} \frac{\partial \omega_{22t|t}}{\partial \boldsymbol{\theta}'} &= \frac{\partial \omega_{22t|t-1}}{\partial \boldsymbol{\theta}'} - 2\omega_{21t|t-1}\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c}\frac{\partial \omega_{21t|t-1}}{\partial \boldsymbol{\theta}'} - 2\omega_{21t|t-1}^2\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\frac{\partial \mathbf{c}}{\partial \boldsymbol{\theta}'} \\ &\quad +[\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta}) \otimes \omega_{21t|t-1}^2\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})]\frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}'}. \end{aligned}$$

However, in order to derive the LM test we only need to evaluate these derivatives under the null of $H_0 : \rho_2 = 0$. In that case,

$$x_{t|t-1} = \rho_1 x_{t-1|t-1}$$

and

$$\frac{\partial x_{t|t-1}}{\partial \boldsymbol{\theta}'} = x_{t-1|t-1}\frac{\partial \rho_1}{\partial \boldsymbol{\theta}'} + x_{t-2|t-1}\frac{\partial \rho_2}{\partial \boldsymbol{\theta}'} + \rho_1 \frac{\partial x_{t-1|t-1}}{\partial \boldsymbol{\theta}'}.$$

Similarly,

$$\boldsymbol{\Omega}_{t|t-1} = \begin{pmatrix} \rho_1^2 \omega_{11t-1|t-1} + 1 & \rho_1 \omega_{11t-1|t-1} \\ \rho_1 \omega_{11t-1|t-1} & \omega_{11t-1|t-1} \end{pmatrix}$$

and

$$\frac{\partial vec(\boldsymbol{\Omega}_{t|t-1})}{\partial \boldsymbol{\theta}'} = \mathbf{D}_2 \left[\begin{array}{c} 2\rho_1 \omega_{11t-1|t-1} \cdot \partial\rho_1/\partial\boldsymbol{\theta}' + 2\rho_1 \omega_{21t-1|t-1} \cdot \partial\rho_2/\partial\boldsymbol{\theta}' + \rho_1^2 \cdot \partial\omega_{11t-1|t-1}/\partial\boldsymbol{\theta}' \\ \omega_{11t-1|t-1} \cdot \partial\rho_1/\partial\boldsymbol{\theta}' + \omega_{21t-1|t-1} \cdot \partial\rho_2/\partial\boldsymbol{\theta}' + \rho_1 \cdot \partial\omega_{11t-1|t-1}/\partial\boldsymbol{\theta}' \\ \partial\omega_{11t-1|t-1}/\partial\boldsymbol{\theta}' \end{array} \right],$$

where $\mathbf{D}_2$ is the duplication matrix of order 2. Hence, when $\rho_2 = 0$ we simply require expressions for $\partial x_{t-1|t-1}/\partial\boldsymbol{\theta}'$ and $\partial\omega_{11t-1|t-1}/\partial\boldsymbol{\theta}'$, which unfortunately we can only obtain recursively on the basis of expressions (A7) and (A8).

We also require expressions for $x_{t-2|t-1}$ and $\omega_{21t-1|t-1}$ under the null, as these quantities are associated to the derivatives with respect to ρ_2. In this sense, it is interesting to obtain the derivatives with respect to this parameter when $\rho_2 = 0$, which will be given by

$$\begin{aligned} \frac{\partial \boldsymbol{\mu}_t(\boldsymbol{\theta})}{\partial \rho_2} &= \mathbf{c}\frac{\partial x_{t|t-1}}{\partial \rho_2}, \\ \frac{\partial vec[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}'} &= (\mathbf{c} \otimes \mathbf{c})\frac{\partial \omega_{11t|t-1}}{\partial \rho_2}, \end{aligned} \tag{A9}$$

with

$$\begin{aligned} \frac{\partial x_{t|t-1}}{\partial \rho_2} &= x_{t-2|t-1} + \rho_1 \frac{\partial x_{t-1|t-1}}{\partial \rho_2}, \\ \frac{\partial \omega_{11t|t-1}}{\partial \rho_2} &= 2\rho_1 \omega_{21t-1|t-1} + \rho_1^2 \frac{\partial \omega_{11t-1|t-1}}{\partial \rho_2}, \end{aligned}$$

$$\begin{aligned} \frac{\partial x_{t|t}}{\partial \rho_2} &= (1 - \omega_{11t|t-1}\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c})\frac{\partial x_{t|t-1}}{\partial \rho_2} + \mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})(\mathbf{y}_t - \mathbf{c}x_{t|t-1})\frac{\partial \omega_{11t|t-1}}{\partial \rho_2} \\ &\quad -\omega_{11t|t-1}\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c}\cdot(\mathbf{y}_t - \mathbf{c}x_{t|t-1})'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c}\frac{\partial \omega_{11t|t-1}}{\partial \rho_2} \\ &= (1 - \omega_{11t|t-1}\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c})\left[\frac{\partial x_{t|t-1}}{\partial \rho_2} + \mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})(\mathbf{y}_t - \mathbf{c}x_{t|t-1})\frac{\partial \omega_{11t|t-1}}{\partial \rho_2}\right] \end{aligned}$$

and

$$\begin{aligned} \frac{\partial \omega_{11t|t}}{\partial \rho_2} &= (1 - 2\omega_{11t|t-1}\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c}+)\frac{\partial \omega_{11t|t-1}}{\partial \rho_2} + \omega_{11t|t-1}^2[\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c}]^2\frac{\partial \omega_{11t|t-1}}{\partial \rho_2}, \\ &= (1 - \omega_{11t|t-1}\mathbf{c}'\boldsymbol{\Sigma}_{t|t-1}^{-1}(\boldsymbol{\theta})\mathbf{c})^2\frac{\partial \omega_{11t|t-1}}{\partial \rho_2} \end{aligned}$$

where we have used (A9).

Interestingly, if we use expressions (A5) and (A6), we can finally write

$$\frac{\partial x_{t|t}}{\partial \rho_2} = \varpi_{t|t}\left(\frac{\partial x_{t|t-1}}{\partial \rho_2} + f_{t|t}\frac{\partial \omega_{11t|t-1}}{\partial \rho_2}\right)$$

and

$$\frac{\partial \omega_{11t|t}}{\partial \rho_2} = \varpi_{t|t}^2\frac{\partial \omega_{11t|t-1}}{\partial \rho_2}.$$

B Spectral scores

As we saw before, the spectral approximation to the log-likelihood function requires the computation of the sample periodogram matrix $\mathbf{I_{yy}}(\lambda_j)$. Expression (2), though, is far from ideal from a computational point of view, and for that reason we make use of the Fast Fourier Transform (FFT). Specifically, given the $T \times N$ original real data matrix $\mathbf{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_t, \ldots, \mathbf{y}_T)'$, the FFT creates the centred and orthogonalised $T \times N$ complex data matrix $\mathbf{Z} = (\mathbf{z}_0, \ldots, \mathbf{z}_j, \ldots, \mathbf{z}_{T-1})'$ by effectively premultiplying $\mathbf{Y} - \ell_T \boldsymbol{\pi}'$ by the $T \times T$ Fourier matrix $\mathbf{W}$. On this basis, we can easily compute $\mathbf{I_{yy}}(\lambda_j)$ as $2\pi \mathbf{z}_j \mathbf{z}_j^*$, where $\mathbf{z}_j^*$ is the complex conjugate transpose of $\mathbf{z}_j$. Hence, the spectral approximation to the log-likelihood function for a non-singular $\mathbf{G_{yy}}(\lambda)$ becomes

$$-\frac{NT}{2}\ln(2\pi) - \frac{1}{2}\sum_{j=0}^{T-1}\ln|\mathbf{G_{yy}}(\lambda_j)| - \frac{2\pi}{2}\sum_{j=0}^{T-1}\mathbf{z}_j^*\mathbf{G}_{\mathbf{yy}}^{-1}(\lambda_j)\mathbf{z}_j,$$

which can be regarded as the log-likelihood function of T independent but heteroskedastic complex Gaussian observations.

But since $\mathbf{z}_j$ does not depend on $\boldsymbol{\pi}$ for $j = 1, \ldots, T-1$ because ℓ_T is proportional to the first column of the orthogonal Fourier matrix and $\mathbf{z}_0 = (\bar{\mathbf{y}}_T - \boldsymbol{\pi})$, where $\bar{\mathbf{y}}_T$ is the sample mean of $\mathbf{y}_t$, it immediately follows that the ML of $\boldsymbol{\pi}$ will be $\bar{\mathbf{y}}_T$. As for the remaining parameters, the score function will be given by:

$$\mathbf{d}_j(\boldsymbol{\theta}) = \frac{1}{2}\frac{\partial vec'\left[\mathbf{G_{yy}}(\lambda_j)\right]}{\partial \boldsymbol{\theta}}\left[\mathbf{G}_{\mathbf{yy}}^{-1}(\lambda_j) \otimes \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda_j)\right] vec\left[2\pi \mathbf{z}_j^c \mathbf{z}_j' - \mathbf{G}_{\mathbf{yy}}'(\lambda_j)\right],$$

where $\mathbf{z}_j^c = \mathbf{z}_j^{*\prime}$ is the complex conjugate of $\mathbf{z}_j$. Given that

$$d\mathbf{G_{yy}}(\lambda) = d\mathbf{c}(e^{-i\lambda})G_{xx}(\lambda)\mathbf{c}'(e^{i\lambda}) + \mathbf{c}(e^{-i\lambda})dG_{xx}(\lambda)\mathbf{c}'(e^{i\lambda}) + \mathbf{c}(e^{-i\lambda})G_{xx}(\lambda)d\mathbf{c}'(e^{i\lambda}) + d\mathbf{G_{uu}}(\lambda)$$

(see Magnus and Neudecker (1988)), it immediately follows that

$$\begin{aligned} dvec\left[\mathbf{G_{yy}}(\lambda)\right] = {} & \left[\mathbf{c}(e^{i\lambda})G_{xx}(\lambda) \otimes \mathbf{I}_N\right] d\mathbf{c}(e^{-i\lambda}) + \left[\mathbf{I}_N \otimes \mathbf{c}(e^{-i\lambda})G_{xx}(\lambda)\right] d\mathbf{c}(e^{i\lambda}) \\ & + \left[\mathbf{c}(e^{i\lambda}) \otimes \mathbf{c}(e^{-i\lambda})\right] dG_{xx}(\lambda) + \mathbf{E}_N dvecd\left[\mathbf{G_{uu}}(\lambda)\right], \end{aligned}$$

where $\mathbf{E}_n$ is the unique $n^2 \times n$ "diagonalisation" matrix which transforms $vec(\mathbf{A})$ into $vecd(\mathbf{A})$ as $vecd(\mathbf{A}) = \mathbf{E}_n' vec(\mathbf{A})$, and $\mathbf{K}_{mn}$ is the commutation matrix of orders m and n (see Magnus (1988)). But

$$\mathbf{c}(e^{-i\lambda}) = \sum_{\ell=-F}^{L} \mathbf{c}_\ell(\boldsymbol{\theta}) e^{-i\ell\lambda}, \tag{B10}$$

so

$$d\mathbf{c}(e^{-i\lambda}) = \sum_{\ell=-F}^{L} d\mathbf{c}_\ell(\boldsymbol{\theta}) e^{-i\ell\lambda}.$$

Consequently, we can write

$$\begin{aligned} dvec\left[\mathbf{G}_{\mathbf{yy}}(\lambda)\right] = & \sum_{\ell=-F}^{L} \left\{ \left[e^{-i\ell\lambda}\mathbf{c}(e^{i\lambda}) G_{xx}(\lambda) \otimes \mathbf{I}_N \right] + \left[\mathbf{I}_N \otimes e^{i\ell\lambda}\mathbf{c}(e^{-i\lambda}) G_{xx}(\lambda) \right] \right\} d\mathbf{c}_\ell(\boldsymbol{\theta}) \\ & + \left[\mathbf{c}(e^{i\lambda}) \otimes \mathbf{c}(e^{-i\lambda}) \right] dG_{xx}(\lambda) + \mathbf{E}_N dvecd\left[\mathbf{G}_{\mathbf{uu}}(\lambda)\right]. \end{aligned}$$

Hence, the Jacobian of $vec\left[\mathbf{G}_{\mathbf{yy}}(\lambda)\right]$ will be

$$\begin{aligned} \frac{\partial vec\left[\mathbf{G}_{\mathbf{yy}}(\lambda)\right]}{\partial \boldsymbol{\theta}'} = & \sum_{\ell=-F}^{L} \left\{ \left[e^{-i\ell\lambda}\mathbf{c}(e^{i\lambda}) G_{xx}(\lambda) \otimes \mathbf{I}_N \right] + \left[\mathbf{I}_N \otimes e^{i\ell\lambda}\mathbf{c}(e^{-i\lambda}) G_{xx}(\lambda) \right] \right\} \frac{\partial \mathbf{c}_\ell}{\partial \boldsymbol{\theta}'} \\ & + \left[\mathbf{c}(e^{i\lambda}) \otimes \mathbf{c}(e^{-i\lambda}) \right] \frac{\partial G_{xx}(\lambda)}{\partial \boldsymbol{\theta}'} + \mathbf{E}_N \frac{\partial vecd\left[\mathbf{G}_{\mathbf{uu}}(\lambda)\right]}{\partial \boldsymbol{\theta}'}. \end{aligned}$$

If we combine this expression with the fact that

$$\begin{aligned} & \left[\mathbf{G}_{\mathbf{yy}}^{-1}(\lambda_j) \otimes \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda_j)\right] vec\left[\mathbf{z}_j^c \mathbf{z}_j' - \mathbf{G}_{\mathbf{yy}}'(\lambda_j)\right] \\ = & \; vec\left[2\pi \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda) \mathbf{z}_j^c \mathbf{z}_j' \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda) - \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda)\right] \end{aligned}$$

and $\mathbf{I}_{\mathbf{yy}}'(\lambda) = \mathbf{z}_j^c \mathbf{z}_j'$ we obtain:

$$\begin{aligned} 2\mathbf{d}_j(\boldsymbol{\theta}) = & \sum_{\ell=-F}^{L} \frac{\partial \mathbf{c}_\ell'}{\partial \boldsymbol{\theta}} \left\{ \begin{array}{c} \left[e^{-i\ell\lambda} G_{xx}(\lambda) \mathbf{c}'(e^{i\lambda}) \otimes \mathbf{I}_N\right] \\ + \left[\mathbf{I}_N \otimes e^{i\ell\lambda} G_{xx}(\lambda) \mathbf{c}'(e^{-i\lambda})\right] \end{array} \right\} vec\left[2\pi \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda) \mathbf{I}_{\mathbf{yy}}'(\lambda) \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda) - \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda)\right] \\ & + \frac{\partial G_{xx}(\lambda)}{\partial \boldsymbol{\theta}} \left[\mathbf{c}'(e^{i\lambda}) \otimes \mathbf{c}'(e^{-i\lambda})\right] vec\left[2\pi \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda) \mathbf{I}_{\mathbf{yy}}'(\lambda) \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda) - \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda)\right] \\ & + \frac{\partial vecd'\left[\mathbf{G}_{\mathbf{uu}}(\lambda)\right]}{\partial \boldsymbol{\theta}} \mathbf{E}_N vec\left[2\pi \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda) \mathbf{I}_{\mathbf{yy}}'(\lambda) \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda) - \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda)\right] \\ = & \sum_{\ell=-F}^{L} \frac{\partial \mathbf{c}_\ell'}{\partial \boldsymbol{\theta}} vec \left[\begin{array}{c} 2\pi \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda) \mathbf{I}_{\mathbf{yy}}'(\lambda) \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda) \mathbf{c}(e^{i\lambda}) G_{xx}(\lambda) e^{-i\ell\lambda} - \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda) \mathbf{c}(e^{i\lambda}) G_{xx}(\lambda) e^{-i\ell\lambda} \\ +2\pi e^{i\ell\lambda} G_{xx}(\lambda) \mathbf{c}'(e^{-i\lambda}) \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda) \mathbf{I}_{\mathbf{yy}}'(\lambda) \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda) - e^{i\ell\lambda} G_{xx}(\lambda) \mathbf{c}'(e^{-i\lambda}) \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda) \end{array} \right] \\ & + \frac{\partial G_{xx}(\lambda)}{\partial \boldsymbol{\theta}} vec\left[2\pi \mathbf{c}'(e^{-i\lambda}) \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda) \mathbf{I}_{\mathbf{yy}}'(\lambda) \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda) \mathbf{c}(e^{i\lambda}) - \mathbf{c}'(e^{-i\lambda}) \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda) \mathbf{c}(e^{i\lambda})\right] \\ & + \frac{\partial vecd'\left[\mathbf{G}_{\mathbf{uu}}(\lambda)\right]}{\partial \boldsymbol{\theta}} \mathbf{E}_N vec\left[2\pi \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda) \mathbf{I}_{\mathbf{yy}}'(\lambda) \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda) - \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda)\right]. \end{aligned}$$

Let us now try to interpret the different components of this expression. The first thing to note is that

$$e^{-i\ell\lambda} vec\left[2\pi \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda) \mathbf{I}_{\mathbf{yy}}'(\lambda) \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda) \mathbf{c}(e^{i\lambda}) G_{xx}(\lambda) - \mathbf{G}_{\mathbf{yy}}'^{-1}(\lambda) \mathbf{c}(e^{i\lambda}) G_{xx}(\lambda)\right]$$

and

$$e^{i\ell\lambda}vec\left[2\pi G_{xx}(\lambda)\mathbf{c}'(e^{-i\lambda})\mathbf{G}'^{-1}_{\mathbf{yy}}(\lambda)\mathbf{I}'_{\mathbf{yy}}(\lambda)\mathbf{G}'^{-1}_{\mathbf{yy}}(\lambda)-G_{xx}(\lambda)\mathbf{c}'(e^{-i\lambda})\mathbf{G}'^{-1}_{\mathbf{yy}}(\lambda)\right]$$

are complex conjugates because the conjugate of a product is the product of the conjugates, so it suffices to analyse one of them.

The transfer function of the Wiener-Kolmogorov smoothed values of the common factor is given by

$$G_{xx}(\lambda)\mathbf{c}'(e^{i\lambda})\mathbf{G}^{-1}_{\mathbf{yy}}(\lambda).$$

As a result, the periodogram and spectral density of the smoothed values of the common factor will be

$$\begin{aligned}
I_{x^Kx^K}(\lambda) &= 2\pi G^2_{xx}(\lambda)\mathbf{c}'(e^{i\lambda})\mathbf{G}^{-1}_{\mathbf{yy}}(\lambda)\mathbf{I}_{\mathbf{yy}}(\lambda)\mathbf{G}^{-1}_{\mathbf{yy}}(\lambda)\mathbf{c}(e^{-i\lambda}),\\
G_{x^Kx^K}(\lambda) &= G^2_{xx}(\lambda)\mathbf{c}'(e^{i\lambda})\mathbf{G}^{-1}_{\mathbf{yy}}(\lambda)\mathbf{c}(e^{-i\lambda}),
\end{aligned}$$

respectively, while the spectral density of its final estimation error $x_t - x^K_{t|\infty}$ is

$$\omega(\lambda) = G_{xx}(\lambda) - G_{x^Kx^K}(\lambda).$$

Similarly, the transfer function of the Wiener-Kolmogorov smoothed values of the specific factors will be

$$\mathbf{G}_{\mathbf{uu}}(\lambda)\mathbf{G}^{-1}_{\mathbf{yy}}(\lambda) = \mathbf{I}_N - \mathbf{c}(e^{-i\lambda})G_{xx}(\lambda)\mathbf{c}'(e^{i\lambda})\mathbf{G}^{-1}_{\mathbf{yy}}(\lambda).$$

As a result, the periodogram and spectral density matrix of the smoothed values of the specific factors are given by

$$\begin{aligned}
\mathbf{I}_{\mathbf{u}^K\mathbf{u}^K}(\lambda) &= 2\pi\mathbf{G}_{\mathbf{uu}}(\lambda)\mathbf{G}^{-1}_{\mathbf{yy}}(\lambda)\mathbf{I}_{\mathbf{yy}}(\lambda)\mathbf{G}^{-1}_{\mathbf{yy}}(\lambda)\mathbf{G}_{\mathbf{uu}}(\lambda),\\
\mathbf{G}_{\mathbf{u}^K\mathbf{u}^K}(\lambda) &= \mathbf{G}_{\mathbf{uu}}(\lambda)\mathbf{G}^{-1}_{\mathbf{yy}}(\lambda)\mathbf{G}_{\mathbf{uu}}(\lambda),
\end{aligned}$$

respectively, while the spectral density of their final estimation errors $\mathbf{u}_t - \mathbf{u}^K_{t|\infty}$ is

$$\mathbf{G}_{\mathbf{uu}}(\lambda) - \mathbf{G}_{\mathbf{u}^K\mathbf{u}^K}(\lambda) = \omega(\lambda)\mathbf{c}(e^{-i\lambda})\mathbf{c}'(e^{i\lambda}).$$

Finally, the co-periodogram and co-spectrum between $x^K_{t|\infty}$ and $\mathbf{u}^K_{t|\infty}$ will be

$$\begin{aligned}
\mathbf{I}_{x^K\mathbf{u}^K}(\lambda) &= 2\pi G_{xx}(\lambda)\mathbf{c}'(e^{i\lambda})\mathbf{G}^{-1}_{\mathbf{yy}}(\lambda)\mathbf{I}_{\mathbf{yy}}(\lambda)\mathbf{G}^{-1}_{\mathbf{yy}}(\lambda)\mathbf{G}_{\mathbf{uu}}(\lambda),\\
\mathbf{G}_{x^K\mathbf{u}^K}(\lambda) &= G_{xx}(\lambda)\mathbf{c}'(e^{i\lambda})\mathbf{G}^{-1}_{\mathbf{yy}}(\lambda)\mathbf{G}_{\mathbf{uu}}(\lambda).
\end{aligned}$$

On this basis, if we further assume that $G_{xx}(\lambda) > 0$ and $\mathbf{G}_{\mathbf{uu}}(\lambda) > \mathbf{0}$ we can write

$$\begin{aligned}
&2\pi\mathbf{G}'^{-1}_{\mathbf{yy}}(\lambda)\mathbf{I}'_{\mathbf{yy}}(\lambda)\mathbf{G}'^{-1}_{\mathbf{yy}}(\lambda)\mathbf{c}(e^{i\lambda})G_{xx}(\lambda)e^{-i\ell\lambda} - \mathbf{G}'^{-1}_{\mathbf{yy}}(\lambda)\mathbf{c}(e^{i\lambda})G_{xx}(\lambda)e^{-i\ell\lambda}\\
&\quad = \mathbf{G}'^{-1}_{\mathbf{uu}}(\lambda)\left[2\pi e^{-i\ell\lambda}\mathbf{I}'_{x^K\mathbf{u}^K}(\lambda) - e^{-i\ell\lambda}\mathbf{G}'_{x^K\mathbf{u}^K}(\lambda)\right],\\
&2\pi\mathbf{c}'(e^{-i\lambda})\mathbf{G}'^{-1}_{\mathbf{yy}}(\lambda)\mathbf{I}'_{\mathbf{yy}}(\lambda)\mathbf{G}'^{-1}_{\mathbf{yy}}(\lambda)\mathbf{c}(e^{i\lambda}) - \mathbf{c}'(e^{-i\lambda})\mathbf{G}'^{-1}_{\mathbf{yy}}(\lambda)\mathbf{c}(e^{i\lambda})\\
&\quad = G^{-2}_{xx}(\lambda)\left[2\pi I_{x^Kx^K}(\lambda) - G_{x^Kx^K}(\lambda)\right]
\end{aligned}$$

and

$$2\pi\mathbf{G}_{\mathbf{yy}}^{\prime-1}(\lambda)\mathbf{I}_{\mathbf{yy}}^{\prime}(\lambda)\mathbf{G}_{\mathbf{yy}}^{\prime-1}(\lambda)-\mathbf{G}_{\mathbf{yy}}^{\prime-1}(\lambda)=\mathbf{G}_{\mathbf{uu}}^{\prime-1}(\lambda)\left[2\pi\mathbf{I}_{\mathbf{u}^K\mathbf{u}^K}^{\prime}(\lambda)-\mathbf{G}_{\mathbf{u}^K\mathbf{u}^K}^{\prime}(\lambda)\right]\mathbf{G}_{\mathbf{uu}}^{\prime-1}(\lambda).$$

Therefore, the component of the score associated to $\mathbf{c}_\ell$ will be the sum across frequencies of terms of the form

$$\mathbf{G}_{\mathbf{uu}}^{\prime-1}(\lambda)\left[2\pi e^{-i\ell\lambda}\mathbf{I}_{x^K\mathbf{u}^K}^{\prime}(\lambda)-e^{-i\ell\lambda}\mathbf{G}_{x^K\mathbf{u}^K}^{\prime}(\lambda)\right]$$

(and their conjugate transposes) which capture the difference between the cross-periodogram and cross-spectrum of $x_{t-\ell}^K$ and u_{it}^K inversely weighted by the spectral density of u_{it}. As a result, we can understand this term as arising from the normal equation in the spectral regression of y_{it} onto $x_{t-\ell}$ but taking into account the unobservability of the regressor.

Similarly, the component of the score associated to the parameters that determine $G_{xx}(\lambda)$ will be the cross-product across frequencies of the product of the derivatives of the spectral density of x_t with the difference between the periodogram and spectrum of x_t^K inversely weighted by the squared spectral density of x_t. In this case, we can interpret this term as arising from a marginal log-likelihood function for x_t that takes into account the unobservability of x_t.

Finally, the component of the score associated to the parameters that determine $G_{u_iu_i}(\lambda)$ will be the cross-product across frequencies of the product of the derivatives of the spectral density of u_{it} with the difference between the periodogram and spectrum of u_{it}^K inversely weighted by the squared spectral density of u_{it}. Once again, we can interpret this term as arising from the conditional log-likelihood function of u_{it} given x_t that takes into account the unobservability of u_{t_i}.

We can then exploit the Woodbury formula

$$\begin{aligned}\mathbf{G}_{\mathbf{yy}}^{-1}(\lambda) &= \mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)-\omega(\lambda)\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda})\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda),\\ \omega(\lambda) &= [G_{xx}^{-1}(\lambda)+\mathbf{c}'(e^{i\lambda t})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda t})]^{-1},\end{aligned}$$

which greatly simplifies the computations (see Sentana (2000)). Specifically, we will have that

$$\begin{aligned}&G_{xx}(\lambda)\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{yy}}^{-1}(\lambda)=G_{xx}(\lambda)\mathbf{c}'(e^{i\lambda})\left[\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)-\omega(\lambda)\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda})\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\right]\\ &=\left[1-\omega(\lambda)\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda})\right]G_{xx}(\lambda)\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)=\omega(\lambda)\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda),\end{aligned}$$

so

$$I_{x^Kx^K}(\lambda)=2\pi\omega^2(\lambda)\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{I}_{\mathbf{yy}}(\lambda)\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda})$$

and

$$G_{x^K x^K}(\lambda) = G_{xx}(\lambda)\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{yy}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda})G_{xx}(\lambda)$$
$$= G_{xx}(\lambda)\mathbf{c}'(e^{i\lambda})\left[\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda) - \omega(\lambda)\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda})\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\right]\mathbf{c}(e^{-i\lambda})G_{xx}(\lambda)$$
$$= \left[1 - \omega(\lambda)\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda})\right]G_{xx}^2(\lambda)\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda}) = \omega(\lambda)G_{xx}(\lambda)\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda}).$$

Similarly,

$$\mathbf{G}_{\mathbf{uu}}(\lambda)\mathbf{G}_{\mathbf{yy}}^{-1}(\lambda) = \mathbf{G}_{\mathbf{uu}}(\lambda)\left[\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda) - \omega(\lambda)\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda})\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\right]$$
$$= \mathbf{I}_N - \omega(\lambda)\mathbf{c}(e^{-i\lambda})\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda) = \mathbf{I}_N - \mathbf{c}(e^{-i\lambda})G_{xx}(\lambda)\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{yy}}^{-1}(\lambda),$$

so

$$\mathbf{I}_{\mathbf{u}^K\mathbf{u}^K}(\lambda) = 2\pi\left[\mathbf{I}_N - \omega(\lambda)\mathbf{c}(e^{-i\lambda})\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\right]\mathbf{I}_{\mathbf{yy}}(\lambda)\left[\mathbf{I}_N - \omega(\lambda)\mathbf{c}(e^{i\lambda})\mathbf{c}'(e^{-i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\right]$$

and

$$\mathbf{G}_{\mathbf{u}^K\mathbf{u}^K}(\lambda) = \mathbf{G}_{\mathbf{uu}}(\lambda)\mathbf{G}_{\mathbf{yy}}^{-1}(\lambda)\mathbf{G}_{\mathbf{uu}}(\lambda) = \mathbf{G}_{\mathbf{uu}}(\lambda)\left[\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda) - \omega(\lambda)\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda})\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\right]\mathbf{G}_{\mathbf{uu}}(\lambda)$$
$$= \mathbf{G}_{\mathbf{uu}}(\lambda) - \omega(\lambda)\mathbf{c}(e^{-i\lambda})\mathbf{c}'(e^{i\lambda}).$$

Finally,

$$\mathbf{I}_{x^K\mathbf{u}^K}(\lambda) = 2\pi\omega(\lambda)\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{I}_{\mathbf{yy}}(\lambda)\left[\mathbf{I}_N - \omega(\lambda)\mathbf{c}(e^{i\lambda})\mathbf{c}'(e^{-i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\right]$$

and

$$\begin{aligned}
\mathbf{G}_{x^K\mathbf{u}^K}(\lambda) &= G_{xx}(\lambda)\mathbf{c}'(e^{i\lambda})\left[\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda) - \omega(\lambda)\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda})\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\right]\mathbf{G}_{\mathbf{uu}}(\lambda) \\
&= G_{xx}(\lambda)\mathbf{c}'(e^{i\lambda})\left[\mathbf{I}_N - \omega(\lambda)\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda})\mathbf{c}'(e^{i\lambda})\right] = \omega(\lambda)\mathbf{c}'(e^{i\lambda}).
\end{aligned}$$

We can then use those expressions to efficiently compute the required expressions. In particular, we will get

$$\begin{aligned}
& G_{xx}(\lambda)\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{yy}}^{-1}(\lambda)\mathbf{I}_{\mathbf{yy}}(\lambda)\mathbf{G}_{\mathbf{yy}}^{-1}(\lambda) - G_{xx}(\lambda)\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{yy}}^{-1}(\lambda) \\
= \;& G_{xx}(\lambda)\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{yy}}^{-1}(\lambda)\mathbf{I}_{\mathbf{yy}}(\lambda)\left[\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda) - \omega(\lambda)\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda})\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\right] - \omega(\lambda)\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda) \\
= \;& G_{xx}(\lambda)\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{yy}}^{-1}(\lambda)\mathbf{I}_{\mathbf{yy}}(\lambda)\left[\mathbf{I}_N - \omega(\lambda)\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda})\mathbf{c}'(e^{i\lambda})\right]\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda) - \omega(\lambda)\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda) \\
= \;& \left[\mathbf{I}_{x^K\mathbf{u}^K}(\lambda) - \omega(\lambda)\mathbf{c}'(e^{i\lambda})\right]\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda),
\end{aligned}$$

$$
\begin{aligned}
& \mathbf{G}_{\mathbf{yy}}^{-1}(\lambda)\mathbf{I}_{\mathbf{yy}}(\lambda)\mathbf{G}_{\mathbf{yy}}^{-1}(\lambda) - \mathbf{G}_{\mathbf{yy}}^{-1}(\lambda) \\
= & \left[\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda) - \omega(\lambda)\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda})\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\right]\mathbf{I}_{\mathbf{yy}}(\lambda)\left[\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda) - \omega(\lambda)\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda})\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\right] \\
& - \left[\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda) - \omega(\lambda)\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda})\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\right] \\
& \mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\left[\mathbf{I}_N - \omega(\lambda)\mathbf{c}(e^{-i\lambda})\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\right]\mathbf{I}_{\mathbf{yy}}(\lambda)\left[\mathbf{I}_N - \omega(\lambda)\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda})\mathbf{c}'(e^{i\lambda})\right]\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda) \\
& -\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\left[\mathbf{G}_{\mathbf{uu}}(\lambda) - \omega(\lambda)\mathbf{c}(e^{-i\lambda})\mathbf{c}'(e^{i\lambda})\right]\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda) \\
= & \mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)\left[\mathbf{I}_{\mathbf{u}^K\mathbf{u}^K}(\lambda) - \mathbf{G}_{\mathbf{u}^K\mathbf{u}^K}(\lambda)\right]\mathbf{G}_{\mathbf{uu}}^{-1}(\lambda)
\end{aligned}
$$

and

$$
\mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{yy}}^{-1}(\lambda)\mathbf{I}_{\mathbf{yy}}(\lambda)\mathbf{G}_{\mathbf{yy}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda}) - \mathbf{c}'(e^{i\lambda})\mathbf{G}_{\mathbf{yy}}^{-1}(\lambda)\mathbf{c}(e^{-i\lambda}) = G_{xx}^{-1}(\lambda)[I_{x^Kx^K}(\lambda) - G_{x^Kx^K}(\lambda)]G_{xx}^{-1}(\lambda).
$$

C Autocorrelation structure of a simple Markov switching model

Let s_t denote a binary Markov chain characterised by the following two parameters

$$
\begin{aligned}
P(s_t &= 0|s_{t-1} = 0) = p, \\
P(s_t &= 1|s_{t-1} = 1) = q.
\end{aligned}
$$

As is well known, the stationary distribution of the chain is characterised by

$$
\pi = P(s_t = 1) = \frac{1-p}{2-p-q}.
$$

It is easy to see that we can then write

$$
(s_t - \pi) = (p+q-1)(s_{t-1} - \pi) + \xi_t, \tag{C11}
$$

where

$$
E(\xi_t|s_{t-1} = 0) = E(\xi_t|s_{t-1} = 1) = 0.
$$

The proof of this statement follows from computing the four possible values that ξ_t can take, and the corresponding probabilities conditional on the relevant value of s_{t-1}. In this sense, tedious algebra shows that ξ_t is equal to

$$
\begin{array}{ccc}
p-1 & \text{when} & s_t = 0, s_{t-1} = 0 \\
p & \text{when} & s_t = 1, s_{t-1} = 0 \\
-q & \text{when} & s_t = 0, s_{t-1} = 1 \\
1-q & \text{when} & s_t = 1, s_{t-1} = 1
\end{array}
$$

Therefore, it follows from (C11) that s_t has the autocorrelation structure of an AR(1) with autoregressive coefficient $p+q-1$.

Now let us define the following process

$$x_t = \mu(s_t) + \varepsilon_t,$$

where

$$\mu(s_t) = \begin{cases} \mu_l & \text{if} \quad s_t = 0 \\ \mu_h & \text{if} \quad s_t = 1 \end{cases}$$

and $\varepsilon_t \sim N(0, \sigma^2)$ independently of the past, present and future values of s_t, as well as of the past values of ε_t.

Given that we can write $\mu(s_t)$ as an affine transformation of s_t (i.e. $\mu(s_t) = \mu_l + (\mu_h - \mu_l)s_t$), it follows that $\mu(s_t)$ also has the autocorrelation structure of an AR(1).

Finally, the results on contemporaneous aggregation of ARMA models imply that x_t, which is the sum of an AR(1) and uncorrelated white noise, will have the autocorrelation structure of an ARMA(1,1). Specifically, given that the autocovariance generating function of an AR(1) with autoregressive coefficient α is

$$\frac{\omega^2}{(1 - \alpha L)(1 - \alpha L^{-1})},$$

where ω^2 is the variance of the innovations, the autocovariance function of the contemporaneously aggregated process will be

$$\frac{\omega^2}{(1 - \alpha L)(1 - \alpha L^{-1})} + \sigma^2 = \frac{\omega^2 + \sigma^2(1 - \alpha L)(1 - \alpha L^{-1})}{(1 - \alpha L)(1 - \alpha L^{-1})} = \frac{\lambda^2(1 - \beta L)(1 - \beta L^{-1})}{(1 - \alpha L)(1 - \alpha L^{-1})},$$

where β and λ^2, which are easily obtained by equating coefficients, correspond to the root of the MA polynomial and variance of the univariate Wold residuals, respectively.

Figure 1: P−value discrepancy plots of dynamic specification tests

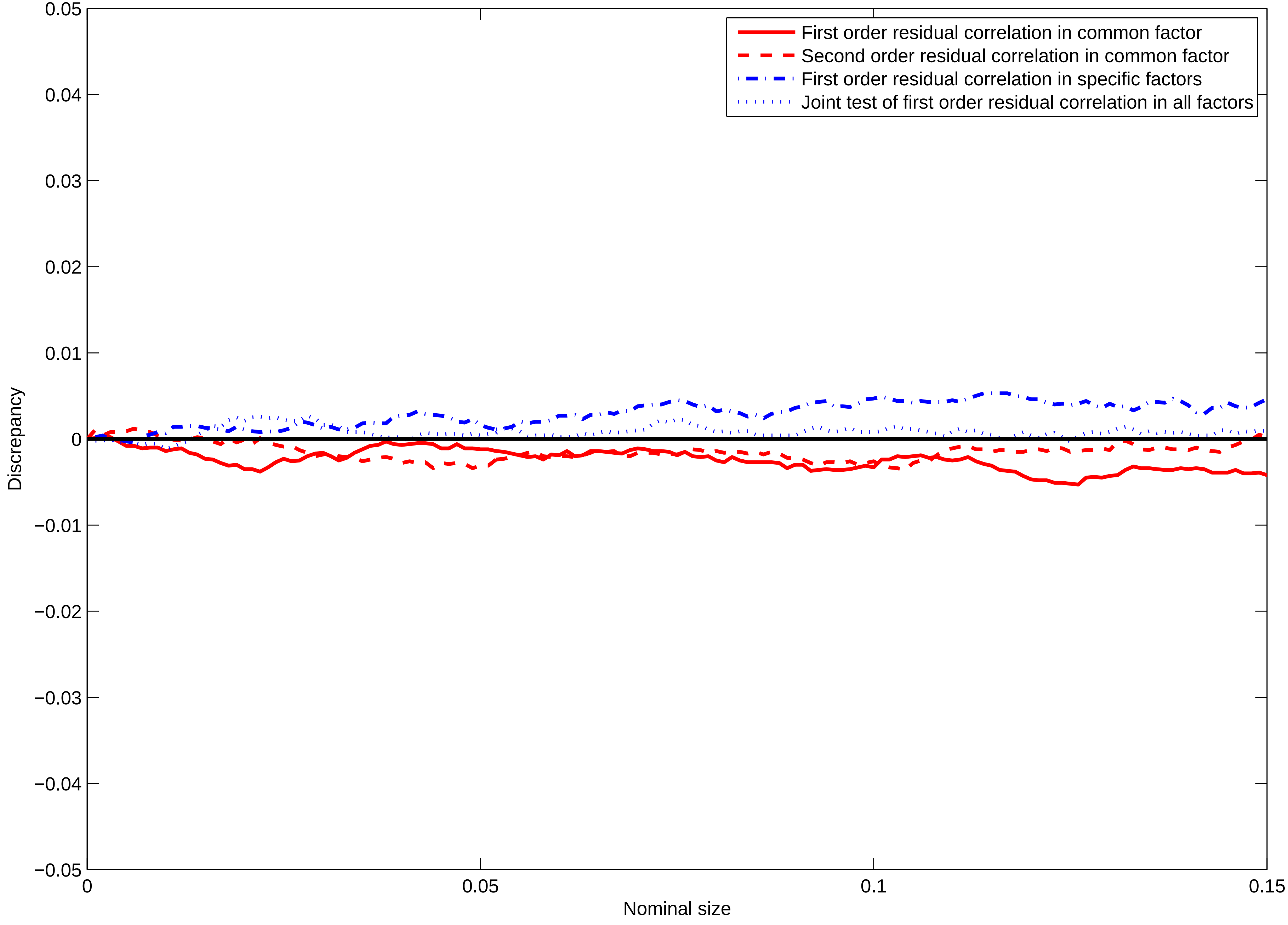


Figure 2: Rejection rates for ARMA(2,1) common factor (ψ=.5)

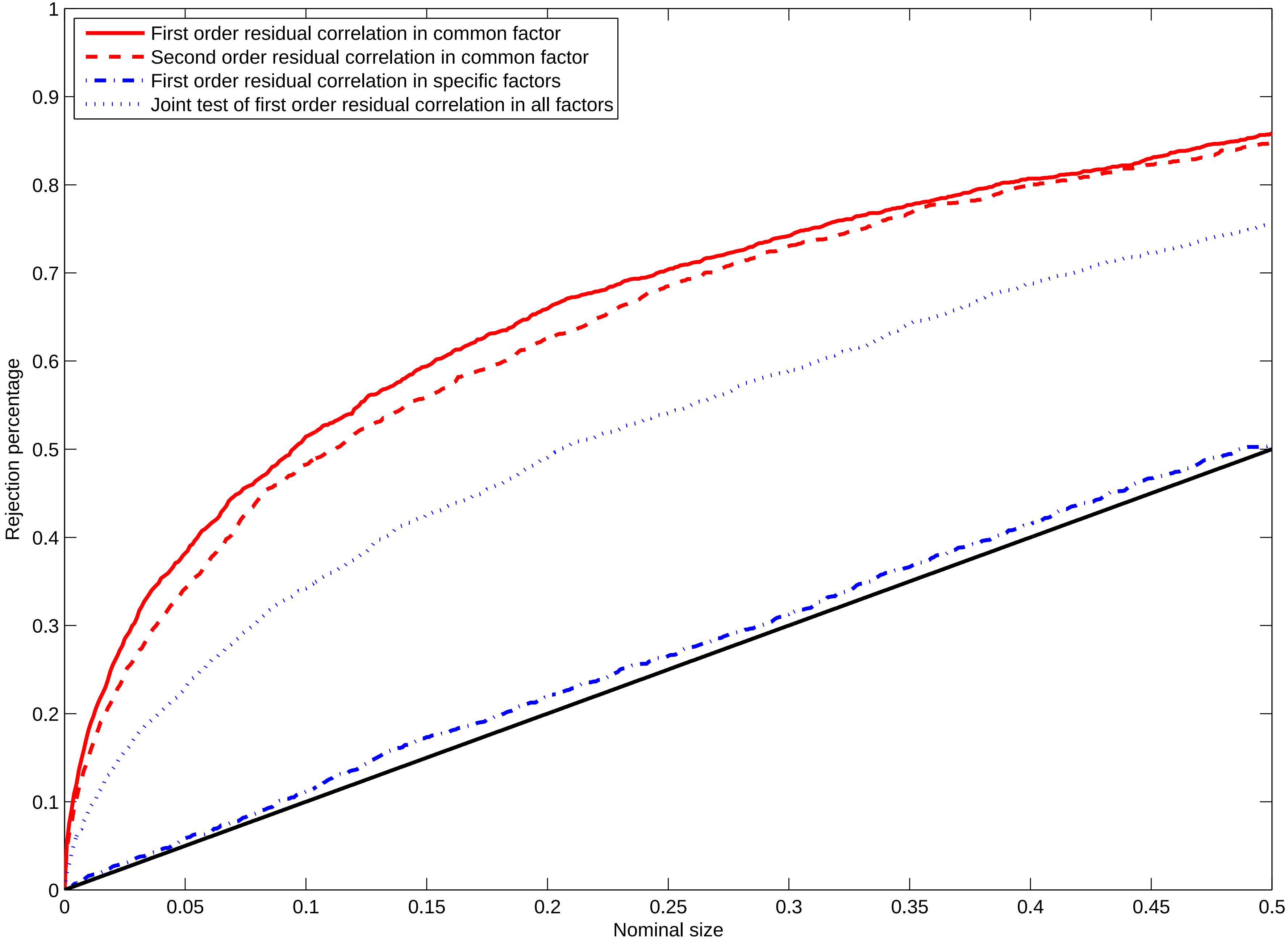


Figure 3: Rejection rates for AR(2) specific factors $(\psi_i = -.2)$

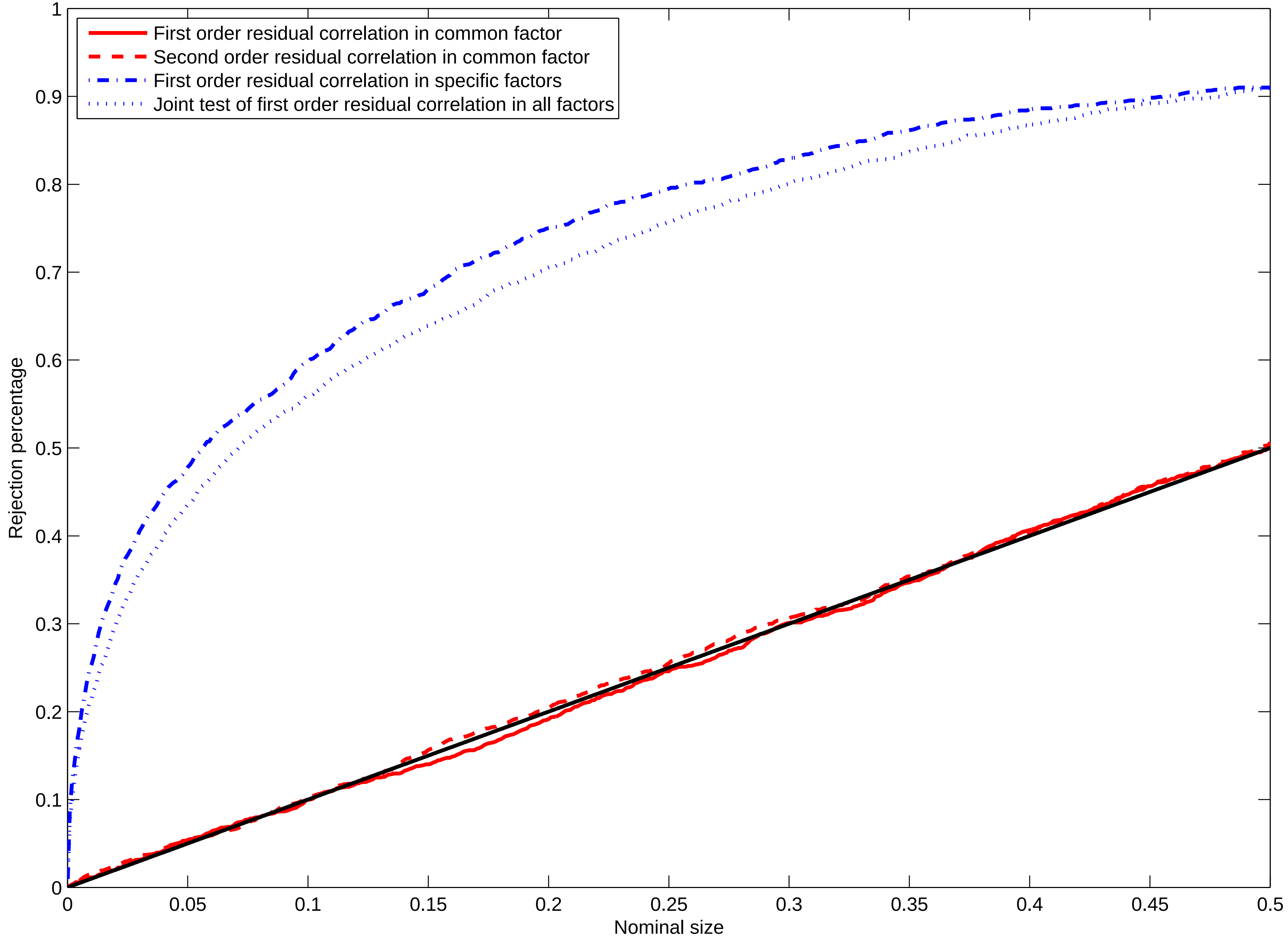


Figure 4: Rejection rates for AR(2) specific factors (ψ_i=−.6)

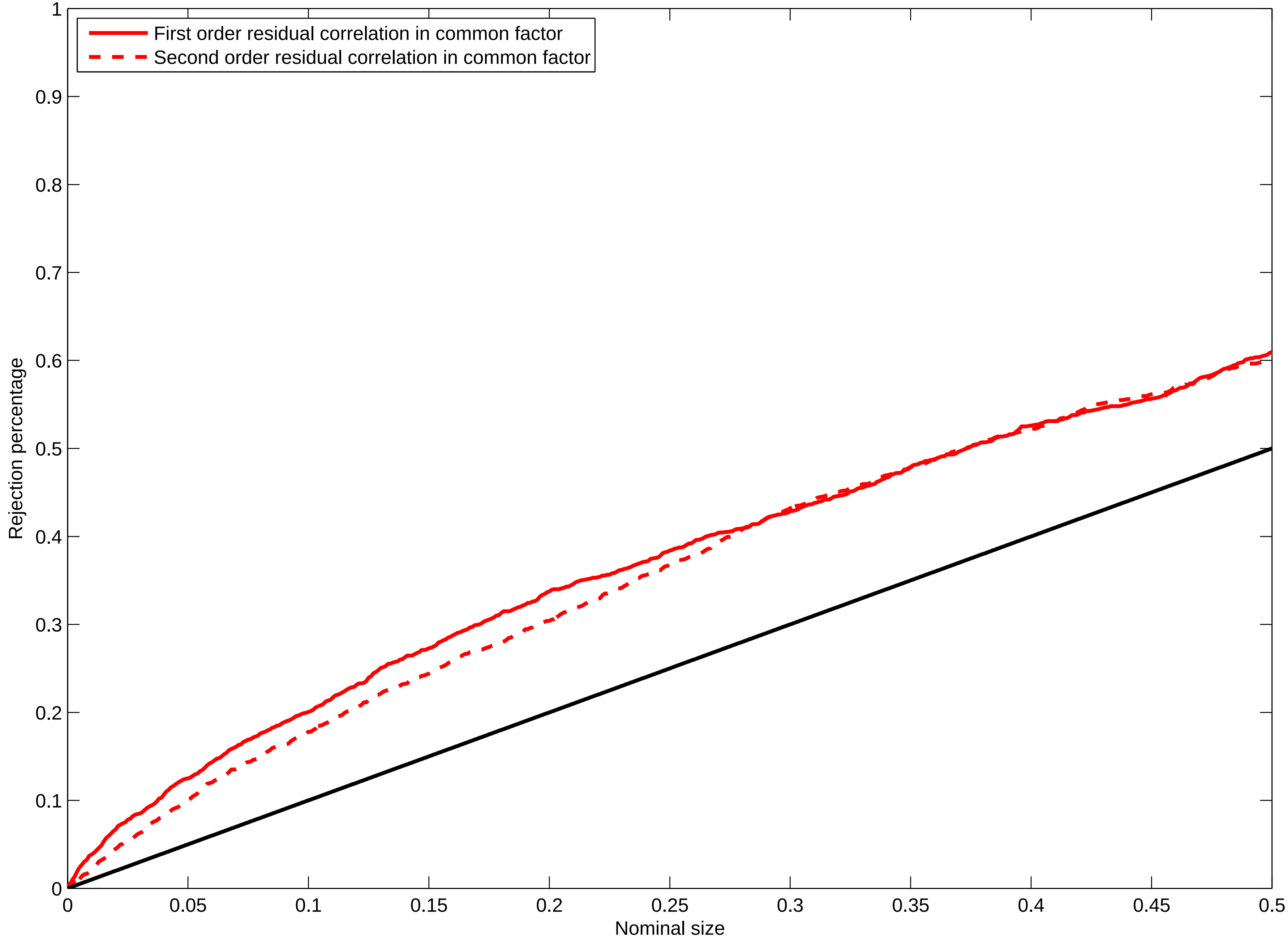
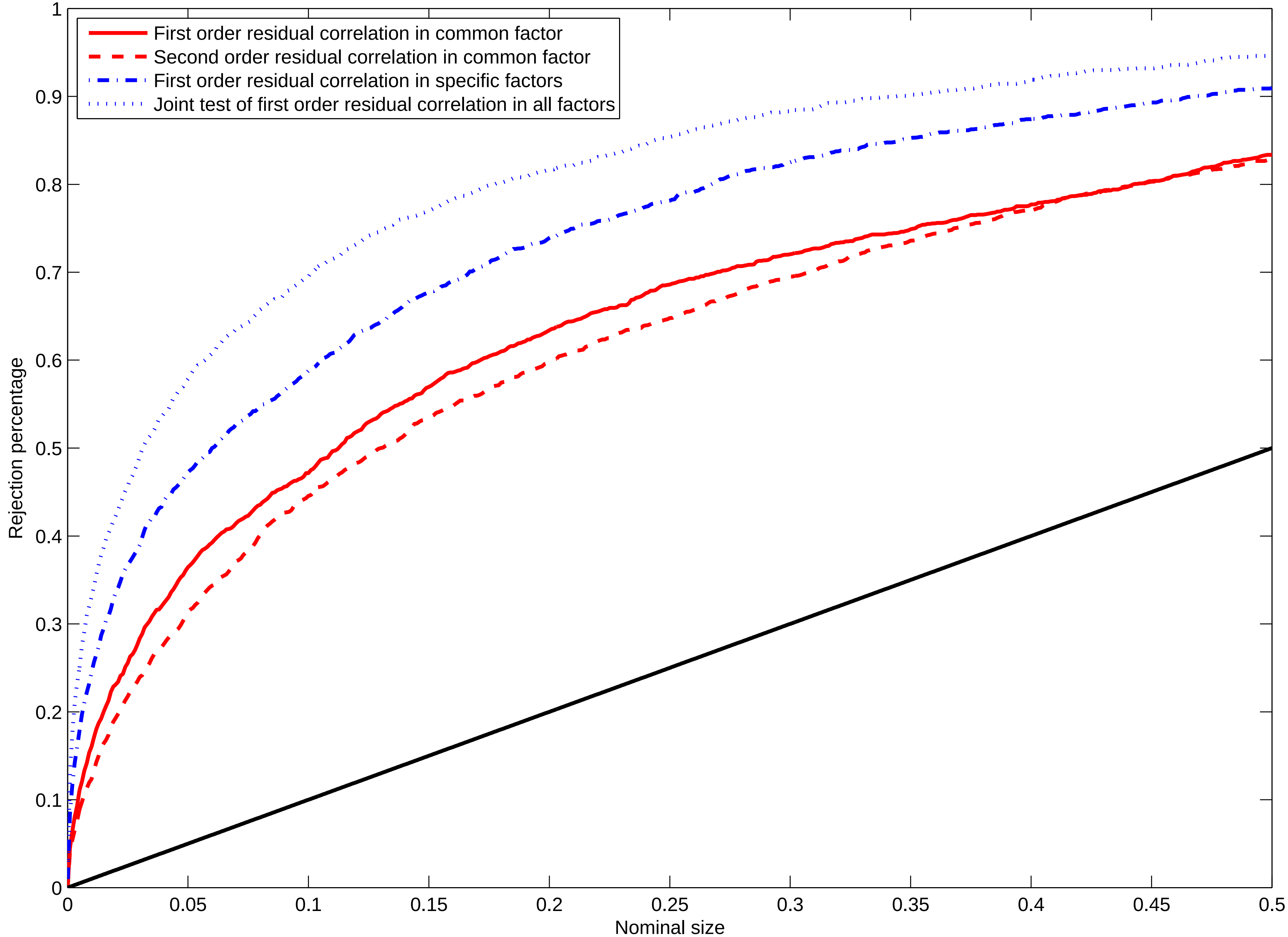


Figure 5: Rejection rates for ARMA(2,1) common factor, AR(2) specific factors (ψ=.5, ψ_i=−.2)



CEMFI WORKING PAPERS

0801 *David Martinez-Miera and Rafael Repullo*: "Does competition reduce the risk of bank failure?".

0802 *Joan Llull*: "The impact of immigration on productivity".

0803 *Cristina López-Mayán:* "Microeconometric analysis of residential water demand".

0804 *Javier Mencía and Enrique Sentana:* "Distributional tests in multivariate dynamic models with Normal and Student *t* innovations".

0805 *Javier Mencía and Enrique Sentana:* "Multivariate location-scale mixtures of normals and mean-variance-skewness portfolio allocation".

0806 *Dante Amengual and Enrique Sentana:* "A comparison of mean-variance efficiency tests".

0807 *Enrique Sentana:* "The econometrics of mean-variance efficiency tests: A survey".

0808 *Anne Layne-Farrar, Gerard Llobet and A. Jorge Padilla:* "Are joint negotiations in standard setting "reasonably necessary"?".

0809 *Rafael Repullo and Javier Suarez:* "The procyclical effects of Basel II".

0810 *Ildefonso Mendez:* "Promoting permanent employment: Lessons from Spain".

0811 *Ildefonso Mendez:* "Intergenerational time transfers and internal migration: Accounting for low spatial mobility in Southern Europe".

0812 *Francisco Maeso and Ildefonso Mendez:* "The role of partnership status and expectations on the emancipation behaviour of Spanish graduates".

0813 *Rubén Hernández-Murillo, Gerard Llobet and Roberto Fuentes:* "Strategic online-banking adoption".

0901 *Max Bruche and Javier Suarez:* "The macroeconomics of money market freezes".

0902 *Max Bruche:* "Bankruptcy codes, liquidation timing, and debt valuation".

0903 *Rafael Repullo, Jesús Saurina and Carlos Trucharte*: "Mitigating the procyclicality of Basel II".

0904 *Manuel Arellano and Stéphane Bonhomme*: "Identifying distributional characteristics in random coefficients panel data models".

0905 *Manuel Arellano, Lars Peter Hansen and Enrique Sentana*: "Underidentification?".

0906 *Stéphane Bonhomme and Ulrich Sauder*: "Accounting for unobservables in comparing selective and comprehensive schooling".

0907 *Roberto Serrano*: "On Watson's non-forcing contracts and renegotiation".

0908 *Roberto Serrano and Rajiv Vohra*: "Multiplicity of mixed equilibria in mechanisms: a unified approach to exact and approximate implementation".

0909 *Roland Pongou and Roberto Serrano*: "A dynamic theory of fidelity networks with an application to the spread of HIV / AIDS".

0910 *Josep Pijoan-Mas and Virginia Sánchez-Marcos:* "Spain is different: Falling trends of inequality".

0911 *Yusuke Kamishiro and Roberto Serrano:* "Equilibrium blocking in large quasilinear economies".

0912 *Gabriele Fiorentini and Enrique Sentana:* "Dynamic specification tests for static factor models".

0913 *Javier Mencía and Enrique Sentana:* “Valuation of VIX derivatives”.

1001 *Gerard Llobet and Javier Suarez:* “Entrepreneurial innovation, patent protection and industry dynamics”.

1002 *Anne Layne-Farrar, Gerard Llobet and A. Jorge Padilla:* “An economic take on patent licensing: Understanding the implications of the “first sale patent exhaustion” doctrine.

1003 *Max Bruche and Gerard Llobet:* “Walking wounded or living dead? Making banks foreclose bad loans”.

1004 *Francisco Peñaranda and Enrique Sentana:* “A Unifying approach to the empirical evaluation of asset pricing models”.

1005 *Javier Suarez:* “The Spanish crisis: Background and policy challenges”.

1006 *Enrique Moral-Benito*: “Panel growth regressions with general predetermined variables: Likelihood-based estimation and Bayesian averaging”.

1007 *Laura Crespo and Pedro Mira:* “Caregiving to elderly parents and employment status of European mature women”.

1008 *Enrique Moral-Benito*: “Model averaging in economics”.

1009 *Samuel Bentolila, Pierre Cahuc, Juan J. Dolado and Thomas Le Barbanchon:* “Two-tier labor markets in the Great Recession: France vs. Spain”.

1010 *Manuel García-Santana and Josep Pijoan-Mas:* “Small Scale Reservation Laws and the misallocation of talent”.

1101 *Javier Díaz-Giménez and Josep Pijoan-Mas:* “Flat tax reforms: Investment expensing and progressivity”.

1102 *Rafael Repullo and Jesús Saurina:* “The countercyclical capital buffer of Basel III: A critical assessment”.

1103 *Luis García-Álvarez and Richard Luger:* “Dynamic correlations, estimation risk, and portfolio management during the financial crisis”.

1104 *Alicia Barroso and Gerard Llobet:* “Advertising and consumer awareness of new, differentiated products”.

1105 *Anatoli Segura and Javier Suarez:* “Dynamic maturity transformation”.

1106 *Samuel Bentolila, Juan J. Dolado and Juan F. Jimeno:* “Reforming an insider-outsider labor market: The Spanish experience”.

1201 *Dante Amengual, Gabriele Fiorentini and Enrique Sentana:* “Sequential estimation of shape parameters in multivariate dynamic models”.

1202 *Rafael Repullo and Javier Suarez:* “The procyclical effects of bank capital regulation”.

1203 *Anne Layne-Farrar, Gerard Llobet and Jorge Padilla:* “Payments and participation: The incentives to join cooperative standard setting efforts”.

1204 *Manuel Garcia-Santana and Roberto Ramos:* “Dissecting the size distribution of establishments across countries”.

1205 *Rafael Repullo:* “Cyclical adjustment of capital requirements: A simple framework”.

1206 *Enzo A. Cerletti and Josep Pijoan-Mas:* “Durable goods, borrowing constraints and consumption insurance”.

1207 *Juan José Ganuza and Fernando Gomez:* “Optional law for firms and consumers: An economic analysis of opting into the Common European Sales Law”.

1208 *Stéphane Bonhomme and Elena Manresa:* “Grouped patterns of heterogeneity in panel data”.

1209 *Stéphane Bonhomme and Laura Hospido:* “The cycle of earnings inequality: Evidence from Spanish Social Security data”.

1210 *Josep Pijoan-Mas and José-Víctor Ríos-Rull:* “Heterogeneity in expected longevities”.

1211 *Gabriele Fiorentini and Enrique Sentana:* “Tests for serial dependence in static, non-Gaussian factor models”.

1301 *Jorge De la Roca and Diego Puga:* “Learning by working in big cities”.

1302 *Monica Martinez-Bravo:* “The role of local officials in new democracies: Evidence from Indonesia”.

1303 *Max Bruche and Anatoli Segura:* “Debt maturity and the liquidity of secondary debt markets”.

1304 *Laura Crespo, Borja López-Noval and Pedro Mira:* “Compulsory schooling, education and mental health: New evidence from SHARELIFE”.

1305 *Lars Peter Hansen:* “Challenges in identifying and measuring systemic risk”.

1306 *Gabriele Fiorentini and Enrique Sentana:* “Dynamic specification tests for dynamic factor models”.