

Weighted Maximum Likelihood for Dynamic Factor Analysis and Forecasting with Mixed-Frequency Data

Francisco Blasques^(a), Siem Jan Koopman^(a,b),
Max Mallee^(a) and Zhaokun Zhang^(a)

(a) VU University Amsterdam and Tinbergen Institute Amsterdam
(b) CREATES, Aarhus University

NBER-NSF Time Series Conference
25,26 September 2015 in Vienna

Motivation

Dynamic factor models are used oftentimes for macroeconomic forecasting.

A key example is forecasting GDP growth.

Within principal components / dynamic factor models, many contributions

- Forni, Hallin, Lippi and Reichlin (RESTAT 2000, JASA 2005)
- Stock and Watson (JASA, JBES 2002)
- Marcellino, Stock and Watson (EER, 2003)
- Doz, Giannone and Reichlin (JEct 2011, RESTAT 2013)
- Bańbura and Rünstler (IJF 2011), Bańbura and Modugno (JAE 2014)
- Jungbacker, Koopman and van der Wel (JEDC 2014),
Jungbacker and Koopman (EctJ 2015),
- Bräuning and Koopman (IJF 2014)

See also the forthcoming Volume 35 of "Advances in Econometrics",

Dynamic Factor Models, 2015, *Eds. E.T. Hillebrand and S.J. Koopman.*

Literature is huge

The previous slide only had references from the 21st century, and then still it is far, far from complete.

This audience, today in Vienna, has many representatives, both from 20th and 21st centuries, but also :

Geweke, Engle, Watson, Tiao, Tsay, Peña, Proietti, Ahn, Reinsel, Velu, West, Boivin, Connor, Quah, Fiorentini, Shumway, Stoffer, Diebold, Sims, Rudebusch, Koop, Korobilis, Ng, Harvey, Frühwirth-Schnatter, Sentana, McCausland, Bernanke, Aguilar, Sargent, McCracken, Bai, Chamberlain, Rothschild, Korajczyk, etc. etc.

So let's conclude, there is a **huge** interest, in many different fields, in

dynamic factor models

What we do

We recognize earlier dynamic factor analysis and forecasting developments while considering the forecasting of GDP growth.

Two issues arise :

- Much effort is devoted to the modelling of so many time series, **big N** , while in the end we only want to forecast **a few key variables**. How should we address this notion to our forecasting model ?
- **Mixed-frequency data** issues are always present in large data sets; they become even more important when the **key variable** has a different frequency.

We discuss both of these issues in this paper.

Our study is related to the paper by Marcellino, Carriero & Clark (2014).

We propose a model-based mixed-frequency dynamic factor state space time series analysis for forecasting and nowcasting.

- Introduction
- Dynamic factor model
- Weighted maximum likelihood estimation
- Monte Carlo study
- Low-frequency representations
- Mixed frequency dynamic factor model
- Illustration : macroeconomic forecasting
- Conclusions and further research

Principal components

Let y_t be the time series of interest, the **key variable**, and let x_t be a very large column vector representing the many "instrumental" variables that are used to improve the forecasting of y_t .

Stock and Watson (2002) advocate to construct principal components series F_t from large data base of x_t variables. Then a parsimonious way to use x_t for the h -steps ahead forecasting of y_t is via the dynamic regression

$$y_{t+h} = \phi(L)y_t + \beta(L)F_t + \epsilon_t,$$

where $\phi(L) = \phi_0 + \phi_1 L + \phi_2 L^2 + \dots$ and $\beta(L) = \beta_0 + \beta_1 L + \beta_2 L^2 + \dots$

Many contributions in the literature has focussed on the appropriate choice of dimension for x_t and, most notably, for F_t .

Many variants of this approach has also appeared in the literature.

Dynamic factor model

The dynamic factor model for the joint analysis of y_t and x_t is given by

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{bmatrix} \Lambda_y \\ \Lambda_x \end{bmatrix} f_t + u_t,$$

where u_t can be assumed to be IID noise but it may also be decomposed into an idiosyncratic dynamic process and IID noise.

The underlying, unobserved vector of dynamic factors f_t can be modelled by the vector autoregressive process

$$f_t = \Phi_1 f_{t-1} + \dots + \Phi_p f_{t-p} + \eta_t,$$

where η_t is typically IID noise, mutually independent of u_t .

The two equations constitute a [linear state space model](#).

Maximum likelihood estimation, quasi-MLE

The number of unknown parameters in the DFM

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{bmatrix} \Lambda_y \\ \Lambda_x \end{bmatrix} f_t + u_t, \quad f_t = \Phi_1 f_{t-1} + \dots + \Phi_p f_{t-p} + \eta_t,$$

is increasing quickly when the dimension of x_t becomes larger and larger.

Some options for maximum likelihood estimation (MLE) :

- Jungbacker and Koopman (2015) : MLE, as done before; direct maximization of loglik wrt all unknown parameters, is feasible with fast loglik evaluation via Kalman filter, after data transformation.
- Doz, Giannone and Reichlin (2011) : two steps – first, replace f_t by F_t and apply regression to both equations; second, replace parameters by these estimates and continue analysis based on Kalman filter.
- Bräuning and Koopman (2014) : replace x_t by F_t and set $\Lambda_x = I$; MLE for remaining coefficients and use this model also for analysis and forecasting:

$$y_t = \Lambda_y f_t + u_{y,t}, \quad F_t = f_t + u_{f,t}.$$

In contributions such as Doz, Giannone and Reichlin (2011, 2013), Bańbura and Rünstler (2011), Bańbura and Modugno (2014), Jungbacker, Koopman and van der Wel (2014), Jungbacker and Koopman (2015) and Bräuning and Koopman (2014), **state space model and Kalman filter** are adopted for estimation, analysis and forecasting.

All estimation procedures above are likelihood-based.

However, dynamic factor model is likely to be **misspecified**... hence we refer to it as quasi-MLE.

But quasi-MLE does not address the different roles of y_t and x_t : y_t being the **key variable** and x_t being the large vector of **instruments**.

DFM and MLE

For the DFM

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{bmatrix} \Lambda_y \\ \Lambda_x \end{bmatrix} f_t + u_t, \quad f_t = \Phi_1 f_{t-1} + \dots + \Phi_p f_{t-p} + \eta_t,$$

we collect all unknown parameters in vector ψ .

The loglikelihood function is given by

$$\mathcal{L}(\psi, f_1) := \log p(y, x; \psi) = \log p(y|x; \psi) + \log p(x; \psi).$$

All series have **equal** importance in this loglikelihood function.

But we are only interested in forecasting y_t accurately...

Instead of maximizing $\ell = p(y, x; \psi) = p(y|x; \psi) \times p(x; \psi)$, perhaps we should maximize

$$\ell(w) = p(y|x; \psi)^w \times p(x; \psi)^{(2-w)}, \quad 1 \leq w < 2.$$

Weighted maximum likelihood estimation

The main idea of the weighted loglikelihood function is to replace

$$\mathcal{L}(\psi, f_1) := \log p(y, x; \psi) = \log p(y|x; \psi) + \log p(x; \psi),$$

by

$$\mathcal{L}_W(\psi, f_1) := W \log p(y|x; \psi) + \log p(x; \psi),$$

with $W > 1$.

The value of W can be pre-fixed or it can be determined by another criterion, for example the minimization of the **out-of-sample MSFE**, (mean squared forecast error), in a **cross-validation** setting.

Note : as W becomes larger, the contribution of x becomes negligible for the estimation of ψ BUT x remains to take full part in the forecasting of y .

Despite this ad-hoc nature, the weighted ML (WML) parameter estimates have the usual asymptotic properties of existence, consistency and asymptotic normality, also when the DFM is misspecified.

Weighted maximum likelihood : asymptotic results

- properties of the weighted maximum likelihood estimator are derived in the paper: for any choice of weight $w := W^{-1} \in [0, 1]$;
- when the model is correctly specified, then the WML estimator $\hat{\psi}_T(w)$ is consistent and asymptotically normal for the true parameter vector $\psi_0 \in \Psi$.
- when the model is misspecified, we show that $\hat{\psi}_T(w)$ is consistent and asymptotically normal for a pseudo-true parameter $\psi_0^*(w) \in \Psi$ that minimizes a transformed Kullback–Leibler (KL) divergence between the true probability measure of the data and the measure implied by the model.
- we show that the transformed KL divergence takes the form of a pseudo-metric that gives more weight to fitting the conditional density of y_t when $W > 1$ or $0 < w < 1$.
- for special case $w = 1$, we obtain the classical pseudo-true parameter $\psi_0^*(1) \in \Psi$ of the ML estimator that minimizes the KL divergence.

Weighted maximum likelihood : Monte Carlo study

DGP 1 for $z_t = (y_t, x_t')'$:

$$z_t = \beta_z f_t + u_t + \varepsilon_t, \quad \varepsilon_\tau \sim NID(0, \sigma_\varepsilon^2 I),$$

where both f_t and u_t are AR(1)'s with $\phi = 0.8$. Factor loadings in β_z for y is unity and for the i th x variable i^{-1} . The variance of the AR(1) disturbances is set to 0.25 and $\sigma_\varepsilon^2 = 0.5$.

DGP 2 for $z_t = (y_t, x_t')'$:

$$z_t = \Phi z_{t-1} + \varepsilon_t, \quad \varepsilon_\tau \sim NID(0, \sigma_\varepsilon^2 I),$$

with diagonal values of Φ equal to 0.80 and off-diagonals are randomly generated $[-0.5, 0.5]$ st z_t is stationary. Diagonal variance matrix for VAR(1) disturbances with variances set to 0.25 and $\sigma_\varepsilon^2 = 0.5$.

Weighted maximum likelihood : Monte Carlo study

Scenario 1 "underspecification" : DGP 1 but we consider DFM that has only common dynamic factors, NOT the idiosyncratic dynamic factors u_t , that is

$$z_t = \beta_z f_t + \varepsilon_t, \quad \varepsilon_\tau \sim NID(0, \sigma_\varepsilon^2).$$

Scenario 2 "misspecification" : DGP 2 but we consider DFM with common dynamic factors only, NOT the idiosyncratic dynamic factors u_t , that is

$$z_t = \beta_z f_t + \varepsilon_t, \quad \varepsilon_\tau \sim NID(0, \sigma_\varepsilon^2).$$

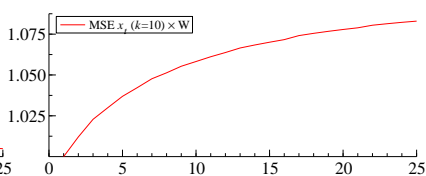
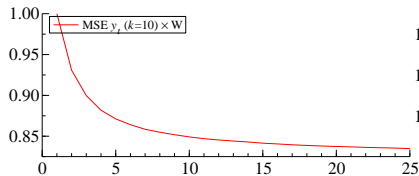
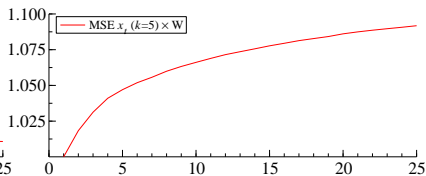
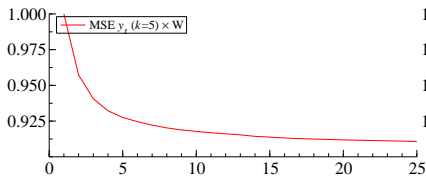
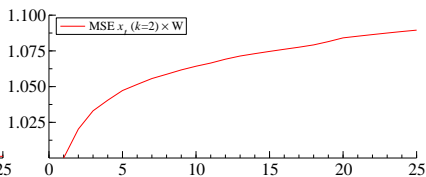
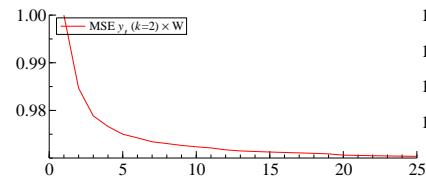
Scenario 3 "correct specification" : DGP 1 and we consider the same model

$$z_t = \beta_z f_t + u_t + \varepsilon_t, \quad \varepsilon_\tau \sim NID(0, \sigma_\varepsilon^2 I).$$

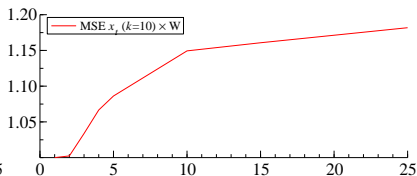
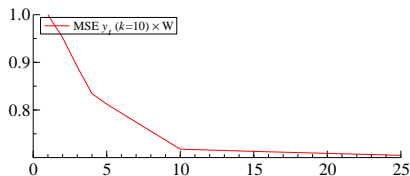
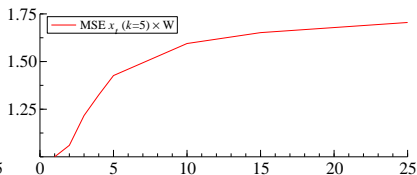
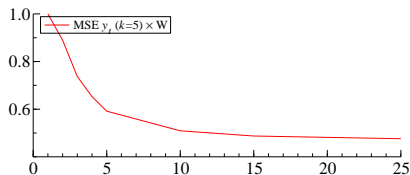
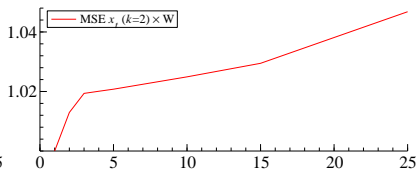
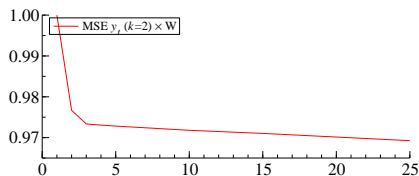
Monte Carlo results : average MSE for y

W	Sc 1 – underspec			Sc 2 – misspec			Sc 3 – c
	$k = 2$	$k = 5$	$k = 10$	$k = 2$	$k = 5$	$k = 10$	$k = 2$
1	1.000	1.000	1.000	1.000	1.000	1.000	1.0000
2	0.983	0.962	0.931	0.977	0.890	0.952	0.9996
3	0.974	0.947	0.889	0.973	0.737	0.891	0.9994
5	0.970	0.938	0.865	0.973	0.592	0.812	0.9992
10	0.968	0.928	0.844	0.972	0.509	0.718	0.9990
25	0.966	0.920	0.831	0.969	0.476	0.705	0.9988
1000	0.965	0.914	0.809	0.965	0.442	0.685	0.9986

Monte Carlo results : Sc 1, average MSE



Monte Carlo results : Sc 2, average MSE



Low-frequency representation

Example is a monthly time series x_τ^m for a variable x that is observed on a monthly basis (m) with monthly time index τ .

The monthly time series can be vectorized into a quarterly process for the 3×1 observed vector x_t^q with quarterly time index t and for

$$x_t^q = \begin{pmatrix} x_{t,1}^q \\ x_{t,2}^q \\ x_{t,3}^q \end{pmatrix} \equiv \begin{pmatrix} x_{3(t-1)+1}^m \\ x_{3(t-1)+2}^m \\ x_{3(t-1)+3}^m \end{pmatrix},$$

where $x_{t,i}^q$ is i -th element of x_t^q and with i being i th month of quarter t .

We further have

$$t = 1, \dots, n, \quad i = 1, 2, 3, \quad \tau = 1, \dots, 3n.$$

We can also represent monthly or quarterly series into yearly vector series.

Low-frequency representation: AR(1)

Consider the monthly AR(1) process

$$\begin{aligned}x_{\tau}^m &= \phi x_{\tau-1}^m + \varepsilon_{\tau}^m \\&= \phi^2 x_{\tau-2}^m + \phi \varepsilon_{\tau-1}^m + \varepsilon_{\tau}^m \\&= \phi^3 x_{\tau-3}^m + \phi^2 \varepsilon_{\tau-2}^m + \phi \varepsilon_{\tau-1}^m + \varepsilon_{\tau}^m,\end{aligned}$$

where $\varepsilon_{\tau}^m \sim NID(0, \sigma_{\varepsilon}^2)$.

The model representation for quarterly vector

$$x_t^q = (x_{3(t-1)+1}^m, x_{3(t-1)+2}^m, x_{3(t-1)+3}^m)'$$

is the VAR(1) process $x_t^q = T x_{t-1}^q + R \varepsilon_t^q$ where

$$T = \begin{pmatrix} 0 & 0 & \phi \\ 0 & 0 & \phi^2 \\ 0 & 0 & \phi^3 \end{pmatrix}, \quad R = \begin{pmatrix} 1 & 0 & 0 \\ \phi & 1 & 0 \\ \phi^2 & \phi & 1 \end{pmatrix},$$

and $\varepsilon_t^q = (\varepsilon_{3(t-1)+1}^m, \varepsilon_{3(t-1)+2}^m, \varepsilon_{3(t-1)+3}^m)'$.

Low-frequency representation: AR(3)

Consider the monthly AR(3) process

$$x_{\tau}^m = \phi_1 x_{\tau-1}^m + \phi_2 x_{\tau-2}^m + \phi_3 x_{\tau-3}^m + \varepsilon_{\tau}^m,$$

where $\varepsilon_{\tau}^m \sim NID(0, \sigma_{\varepsilon}^2)$.

Then model representation for quarterly vector x_t^q is the VAR(1) process $x_t^q = T x_{t-1}^q + R \varepsilon_t^q$ where

$$T = \begin{pmatrix} \phi_3 & \phi_2 & \phi_1 \\ \phi_1 \phi_3 & \phi_1 \phi_2 + \phi_3 & \phi_1^2 + \phi_2 \\ \phi_1^2 \phi_3 + \phi_2 \phi_3 & \phi_1^2 \phi_2 + \phi_1 \phi_3 + \phi_2^2 & \phi_1^3 + 2\phi_1 \phi_2 + \phi_3 \end{pmatrix},$$

$$R = \begin{pmatrix} 1 & 0 & 0 \\ \phi_1 & 1 & 0 \\ \phi_1^2 + \phi_2 & \phi_1 & 1 \end{pmatrix},$$

and $\varepsilon_t^q = (\varepsilon_{3(t-1)+1}^m, \varepsilon_{3(t-1)+2}^m, \varepsilon_{3(t-1)+3}^m)'$.

Low-frequency representation: $AR(p)$

Similar representations are available for the monthly $AR(p)$ process

$$x_t^m = \phi_1 x_{t-1}^m + \phi_2 x_{t-2}^m + \dots + \phi_p x_{t-p}^m + \varepsilon_t^m,$$

where $\varepsilon_t^m \sim NID(0, \sigma_\varepsilon^2)$.

For $p > 3$, we require the linear state space representation $x_t^q = Z\alpha_t + H\varepsilon_t$ and $\alpha_{t+1} = T\alpha_t + R\eta_t$ where dimension of state vector is $p \times 1$.

These state space representations are straightforward and not used in econometrics... but these representations are known and used in the engineering and time series literature.

However, it turns out that exact likelihood evaluation for monthly AR models with larger p , is also computed **faster** when AR process is represented in quarterly state space representation !

Note : for quarterly series, less frequent updating necessary than for monthly series !

Low-frequency updating is computationally more efficient

p	Computing times (in seconds)			State dim		
	Monthly ($n = 12K$)	Quarterly ($n = 4K$)	Yearly ($n = 1K$)	M	Q	Y
1	10	13	61	1	3	12
2	11	16	67	2	3	12
3	26	18	76	3	3	12
4	41	27	85	4	4	12
5	59	40	92	5	5	12
6	83	56	100	6	6	12
7	106	73	108	7	7	12
8	129	90	116	8	8	12
9	154	111	124	9	9	12
10	191	137	133	10	10	12
11	226	162	139	11	11	12
12	265	190	146	12	12	12

Mixed-frequency – dynamic factor model – forecasting/nowcasting

Much work is done on these topics :

- Bridge models : Baffigi, Golinelli and Parigi (2004)
- MIDAS : Ghysels, Foroni, Marcellino and Schumacher (2012)
- MF-DFM : Mariano & Murasawa (2004), Marcellino, Carriero and Clark (2014) Aruoba, Diebold & Scotti (2008), etc.
- Forecasting/nowcasting : Bańbura et al. (2013), Bräuning and Koopman (2014), Hindrayanto et al. (2014)

Mixed-Frequency Model in Low-Frequency representation

Consider

- monthly observed variable x_τ^m is modeled as the AR(1) process
 $x_{\tau+1}^m = \phi_x x_\tau^m + \varepsilon_\tau^m$;
- quarterly observed variable y_t is modeled by the AR(1) process
 $y_{t+1} = \phi_y y_t + \xi_t$;

We combine the two series into a quarterly vector process

$$\begin{pmatrix} y_{t+1} \\ x_{t+1,1}^q \\ x_{t+1,2}^q \\ x_{t+1,3}^q \end{pmatrix} = \begin{pmatrix} \phi_y & 0 & 0 & 0 \\ 0 & 0 & 0 & \phi_x \\ 0 & 0 & 0 & \phi_x^2 \\ 0 & 0 & 0 & \phi_x^3 \end{pmatrix} \begin{pmatrix} y_t \\ x_{t,1}^q \\ x_{t,2}^q \\ x_{t,3}^q \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \phi_x & 1 & 0 \\ 0 & \phi_x^2 & \phi_x & 1 \end{pmatrix} \begin{pmatrix} \xi_t \\ \varepsilon_{t,1}^q \\ \varepsilon_{t,2}^q \\ \varepsilon_{t,3}^q \end{pmatrix},$$

Here the time series are seemingly unrelated.

Mixed Frequency Dynamic Factor Model: matrix notation

In matrix notation, we have

$$\begin{pmatrix} y_t \\ x_{t,1}^q \\ x_{t,2}^q \\ x_{t,3}^q \end{pmatrix} = \begin{pmatrix} \beta_y & \beta_y & \beta_y \\ \beta_x & 0 & 0 \\ 0 & \beta_x & 0 \\ 0 & 0 & \beta_x \end{pmatrix} \begin{pmatrix} f_{t,1}^q \\ f_{t,2}^q \\ f_{t,3}^q \end{pmatrix} + \begin{pmatrix} \xi_t \\ \varepsilon_{t,1}^q \\ \varepsilon_{t,2}^q \\ \varepsilon_{t,3}^q \end{pmatrix},$$

with the vector autoregressive process for f_t^q given by

$$f_{t+1}^q = T_f f_t^q + R_f \eta_t^q,$$

It is straightforward to generalize the model further:

- loading matrix structure;
- higher or lower lag loadings on monthly factor;
- dynamic specification for monthly factor;
- covariance structure for disturbances.

High-frequency representation with "missing" values

Mariano and Murasawa (2003) approach : treat all series in high-frequency (monthly); insert missings for low-frequency (quarterly), that is

$$\begin{bmatrix} \cdot & \cdot & y_3 & \cdot & \cdot & y_6 & \cdot & \dots & y_{3n} \\ x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & \dots & x_{3n} \end{bmatrix},$$

with model

$$\begin{pmatrix} \tilde{y}_\tau^m \\ x_\tau \end{pmatrix} = \begin{pmatrix} \beta_y g(f_\tau) \\ \beta_x f_\tau \end{pmatrix} + \begin{pmatrix} \xi_\tau \\ \varepsilon_\tau \end{pmatrix}$$

where

$$g(a_\tau) = \frac{1}{3}a_\tau + \frac{2}{3}a_{\tau-1} + a_{\tau-2} + \frac{2}{3}a_{\tau-3} + \frac{1}{3}a_{\tau-4}.$$

and factor f_τ follows an AR process.

Low-frequency solution by averaging monthly into quarterly

We average the monthly series into a quarterly series and we model the low frequency only.

We have

$$\begin{pmatrix} y_t \\ \bar{x}_t \end{pmatrix} = \begin{bmatrix} \beta_y \\ \beta_x \end{bmatrix} \begin{pmatrix} f_t \end{pmatrix} + \begin{pmatrix} \xi_t \\ \bar{\varepsilon}_t \end{pmatrix},$$

Mixed Frequency Dynamic Factor Model : monthly factors

In matrix notation, MF dynamic factor model with k monthly variables and with a monthly factor is given by

$$\begin{pmatrix} y_t \\ x_{t,1}^{q,(1)} \\ x_{t,2}^{q,(1)} \\ x_{t,3}^{q,(1)} \\ \vdots \\ x_{t,1}^{q,(k)} \\ x_{t,2}^{q,(k)} \\ x_{t,3}^{q,(k)} \end{pmatrix} = \begin{pmatrix} \beta_y & \beta_y & \beta_y \\ \beta_x^{(1)} & 0 & 0 \\ 0 & \beta_x^{(1)} & 0 \\ 0 & 0 & \beta_x^{(1)} \\ \vdots & \vdots & \vdots \\ \beta_x^{(k)} & 0 & 0 \\ 0 & \beta_x^{(k)} & 0 \\ 0 & 0 & \beta_x^{(k)} \end{pmatrix} \begin{pmatrix} f_{t,1}^q \\ f_{t,2}^q \\ f_{t,3}^q \end{pmatrix} + \begin{pmatrix} \xi_t \\ \varepsilon_{t,1}^{q,(1)} \\ \varepsilon_{t,2}^{q,(1)} \\ \varepsilon_{t,3}^{q,(1)} \\ \vdots \\ \varepsilon_{t,1}^{q,(k)} \\ \varepsilon_{t,2}^{q,(k)} \\ \varepsilon_{t,3}^{q,(k)} \end{pmatrix},$$

with the vector autoregressive process for f_t^q given by

$$f_{t+1}^q = T_f f_t^q + R_f \eta_t^q,$$

representing monthly dynamics for the factors.

Mixed Frequency Dynamic Factor Model : quarterly factors

In matrix notation, MF dynamic factor model with k monthly variables and with a quarterly factor is given by

$$\begin{pmatrix} y_t \\ x_{t,1}^{q,(1)} \\ x_{t,2}^{q,(1)} \\ x_{t,3}^{q,(1)} \\ \vdots \\ x_{t,1}^{q,(k)} \\ x_{t,2}^{q,(k)} \\ x_{t,3}^{q,(k)} \end{pmatrix} = \begin{pmatrix} \beta_y \\ \beta_x^{(1)} \\ \beta_x^{(1)} \\ \beta_x^{(1)} \\ \vdots \\ \beta_x^{(k)} \\ \beta_x^{(k)} \\ \beta_x^{(k)} \end{pmatrix} f_t + \begin{pmatrix} \xi_t \\ \varepsilon_{t,1}^{q,(1)} \\ \varepsilon_{t,2}^{q,(1)} \\ \varepsilon_{t,3}^{q,(1)} \\ \vdots \\ \varepsilon_{t,1}^{q,(k)} \\ \varepsilon_{t,2}^{q,(k)} \\ \varepsilon_{t,3}^{q,(k)} \end{pmatrix},$$

with an quarterly autoregressive process for f_t .

Empirical study: revisiting Mariano & Murasawa (2003)

Indicator	Description
	Quarterly
GDP	Real GDP (billions of chained 1996 \$, SA, AR)
	Monthly
EMP	Employees on non-agricultural payrolls (thousands, SA)
INC	Personal income less transf.paym (bns chained 1996 \$, SA, AR)
IIP	Index of industrial production (1992 = 100, SA)
SLS	Manufacturing and trade sales (mns chained \$, SA)

Original data set : January 1959 upto December 2000.

Extended data set : January 1960 upto December 2009.

Parameter estimates MFI, MM data

Parameter	MFI Model				
	$\Delta \ln \text{GDP}$	$\Delta \ln \text{EMP}$	$\Delta \ln \text{INC}$	$\Delta \ln \text{IIP}$	$\Delta \ln \text{SLS}$
β	1.00	0.49 (0.04)	0.81 (0.06)	2.14 (0.13)	1.74 (0.11)
ϕ_F			0.56 (0.05)		
σ_F^2			0.08 (0.01)		
$\phi_{u,1}$	-0.04 (0.08)	0.10 (0.04)	-0.05 (0.04)	-0.05 (0.07)	-0.41 (0.05)
$\phi_{u,2}$	-0.83 (0.07)	0.45 (0.05)	0.03 (0.05)	-0.06 (0.06)	-0.20 (0.05)
$\sigma_{u,2}^2$	0.19 (0.04)	0.02 (0.00)	0.09 (0.01)	0.25 (0.02)	0.61 (0.04)

Parameter estimates MFS-M, MM data

MFS-M Model					
Parameter	$\Delta \ln \text{GDP}$	$\Delta \ln \text{EMP}$	$\Delta \ln \text{INC}$	$\Delta \ln \text{IIP}$	$\Delta \ln \text{SLS}$
β	1.00	0.57 (0.04)	0.90 (0.06)	2.30 (0.13)	1.83 (0.12)
ϕ_F			0.59 (0.04)		
σ_F^2			0.06 (0.01)		
$\phi_{u,1}$	-0.40 (0.09)	0.07 (0.05)	-0.08 (0.05)	-0.01 (0.05)	-0.38 (0.05)
$\phi_{u,2}$	-0.21 (0.16)	0.43 (0.06)	0.01 (0.07)	-0.05 (0.07)	-0.17 (0.07)
$\sigma_{u,2}^2$	0.27 (0.04)	0.02 (0.00)	0.09 (0.01)	0.27 (0.03)	0.64 (0.05)

Parameter estimates MFS-Q, MM data

MFS-Q Model					
Parameter	$\Delta \ln \text{GDP}$	$\Delta \ln \text{EMP}$	$\Delta \ln \text{INC}$	$\Delta \ln \text{IIP}$	$\Delta \ln \text{SLS}$
β	1.00	0.25 (0.02)	0.33 (0.02)	0.72 (0.04)	0.60 (0.04)
ϕ_F			0.69 (0.06)		
σ_F^2			0.25 (0.04)		
$\phi_{u,1}$	-0.30 (0.09)	0.11 (0.05)	0.10 (0.04)	-0.10 (0.05)	-0.37 (0.04)
$\phi_{u,2}$	-0.13 (0.13)	0.24 (0.07)	-0.06 (0.05)	-0.11 (0.06)	-0.20 (0.06)
$\sigma_{u,2}^2$	0.24 (0.03)	0.03 (0.00)	0.10 (0.01)	0.34 (0.02)	0.74 (0.05)

Parameter estimates MFA, MM data

Parameter	MFA Model				
	$\Delta \ln \text{GDP}$	$\Delta \ln \text{EMP}$	$\Delta \ln \text{INC}$	$\Delta \ln \text{IIP}$	$\Delta \ln \text{SLS}$
β	1.00	0.67 (0.06)	0.95 (0.08)	2.18 (0.12)	1.77 (0.11)
ϕ_F			0.68 (0.06)		
σ_F^2			0.26 (0.04)		
$\phi_{u,1}$	-0.27 (0.09)	0.69 (0.11)	-0.05 (0.08)	-0.14 (0.11)	-0.22 (0.08)
$\phi_{u,2}$	-0.11 (0.12)	0.09 (0.11)	-0.03 (0.10)	-0.05 (0.14)	-0.19 (0.10)
$\sigma_{u,2}^2$	0.25 (0.03)	0.06 (0.01)	0.40 (0.05)	0.56 (0.10)	1.11 (0.13)

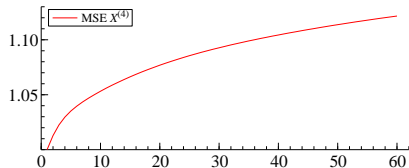
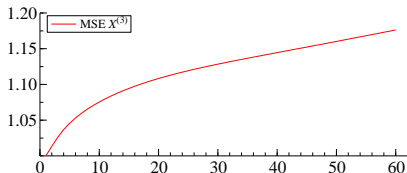
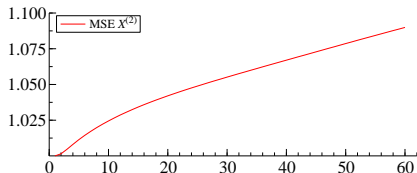
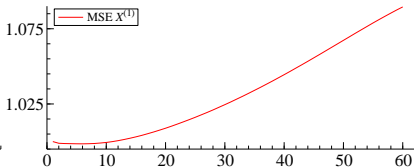
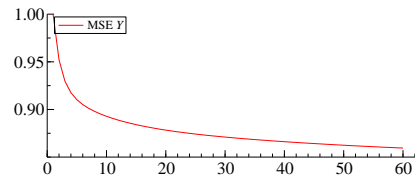
Forecast comparison for US GDP growth 2000-2009

We compare the forecasts of the different mixed-frequency dynamic factor model with the benchmark models "Bridge model" and "MIDAS regressions". Parameter estimates obtained by ML (unweighted).

	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 6$
MFI	0.1779	0.1918	0.2340	0.3156	0.4023
MFS-M	0.1666	0.1730	0.2108	0.2935	0.3986
MFS-Q	0.1765	0.1909	0.2411	0.2989	0.3701
MFA	0.1693			0.2809	0.3754
BM	0.1833	0.2056	0.2455	0.3046	0.4180
MIDAS	0.1597	0.1658	0.2464	0.3635	0.4873

In-sample accuracy using WML

MSEs of in-sample one-step ahead predictions for different W .



Forecast comparison for US GDP with WML

W	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 6$
1	0.1666	0.1730	0.2108	0.2935	0.3917
2	0.1600	0.1689	0.2049	0.2826	0.3708
3	0.1571	0.1674	0.2028	0.2783	0.3614
4	0.1556	0.1670	0.2013	0.2759	0.3534
5	0.1517	0.1703	0.2004	0.2745	0.3662
6	0.1513	0.1560	0.1914	0.2777	0.3733
7	0.1611	0.1668	0.2034	0.2773	0.3715
8	0.1608	0.1670	0.2033	0.2772	0.3699
9	0.1612	0.1682	0.2032	0.2775	0.3683
10	0.1614	0.1690	0.2033	0.2781	0.3662
11	0.1615	0.1698	0.2035	0.2786	0.3577
12	0.1617	0.1705	0.2037	0.2792	0.3572

Forecast comparison for US GDP with WML $W = 6$

	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 6$
MFI	0.1787	0.1885	0.2078	0.2841	0.3629
MFS-M	0.1513	0.1560	0.1914	0.2777	0.3733
MFS-Q	0.1630	0.1676	0.2249	0.2849	0.3670
MFA	0.1576			0.2809	0.3677
BM	0.1833	0.2056	0.2455	0.3046	0.4197
MIDAS	0.1597	0.1658	0.2464	0.3635	0.4873

Forecast comparison for US GDP with optimal WML

	Mean Square Error				
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 6$
MFI	0.1687	0.1765	0.1966	0.2835	0.3559
MFS-M	0.1513	0.1560	0.1914	0.2745	0.3593
MFS-Q	0.1629	0.1670	0.2215	0.2835	0.3621
MFA	0.1576			0.2769	0.3566
BM	0.1833	0.2056	0.2455	0.3046	0.4197
MIDAS	0.1597	0.1658	0.2464	0.3635	0.4873

	Optimal vale of W				
	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 6$
MFI	2	2	2	5	2
MFS-M	6	6	6	5	4
MFS-Q	7	8	3	3	2
MFA	6			2	8

Conclusions

We have presented some further developments for the forecasting of macroeconomic variables using mixed-frequency dynamic factor models:

- estimate parameters by weighted maximum likelihood method;
- base analysis on low-frequency representations of high-frequency dynamics.

More further work can be considered:

- use of low-frequency representations in other mixed-frequency dynamic models;
- carry out a more in-depth study into what type of misspecification can be treated effectively by WML;
- obtain more specific asymptotic results and analysis (under which conditions do we obtain higher asymptotic precision with WML).