Non-stationary Dynamic Factor Models: Methods and applications

Francisco Corona
Instituto Nacional de Estadística y Geografía
Dirección General Adjunta de Investigación

Workshop on Econometrics and Data Science (WEDS)

November 9, 2017.

Motivation

- Dynamic Factor Models (DFMs) are useful for representing the dynamics of a group of N correlated time series through a small number of underlying common factors.
- Nowadays in economics, the availability of large amount of information collected during extended periods, T, grants the opportunity to understand several economic phenomena in a better way.

Motivation

- In economics, large DFMs are useful to adequately modelling the dynamics of a large system of time series, mainly with the following two objectives:
  - Forecasting macroeconomic variables: Stock and Watson (2002a,b), Marcellino *et al.* (2003), Schumacher (2005), Boivin and Ng (2006), etc.
  - Macroeconomic policy making: Bernanke and Boivin (2003), Artis *et al.* (2004), Lahiri and Yao (2004), Bernanke *et al.* (2005), Favero *et al.* (2005), Stock and Watson (2005), etc.

Motivation

- A distinctive feature of macroeconomic time series is non-stationarity which entails the possibility of cointegration in multidimensional systems. Furthermore, cointegration may be regarded as a representation in terms of common factors.
- This analogy suggests that the use of DFMs is, at least conceptually, not restricted to stationary time series, and that it may prove fruitful when dealing with non-stationary vector time series.

Motivation

- However, the most popular way of dealing with large systems of non-stationary macroeconomic time series is by differencing the variables in a univariate fashion.
- The main reason for this extended practice is that the factors are most commonly estimated by Principal Components (PC), a technique which yields a robust asymptotic theory for this case (Bai, 2003).

Motivation

- In this talk, we present the effects of differencing non-stationary DFMs when determining the number of common factors and estimating the factor space.
- We show two novel applications, first, when forecasting the Mexican economic activity using common trends and second, when we study the dynamic structure in the performance behavior from teams in Liga MX.

Determining the number of factors

Introduction

- Cointegration can be related with common factors (common trends), i.e., is clear that if y_{1t} and y_{2t} are cointegrated, they share a common stochastic trend, F_t which is also $I(1)$

$$y_{1t} = \gamma_0 + \gamma_1 F_t + \varepsilon_{1t},$$
$$y_{2t} = \delta_0 + \delta_1 F_t + \varepsilon_{2t},$$

where ε_{1t} and ε_{2t} are the $I(0)$ idiosyncratic components with zero mean and can be weak serial (and cross-section) correlated.

- In this case, given that y_{1t} and y_{2t} are $I(1)$, the stationary linear combination is

$$\delta_1 y_{1t} - \gamma_1 y_{2t} \sim I(0)$$
$$\delta_1\gamma_0 + \delta_1\gamma_1 F_t + \delta_1\varepsilon_{1t} - \delta_0\gamma_1 - \delta_1\gamma_1 F_t - \gamma_1\varepsilon_{2t}$$
$$\approx a + b'\varepsilon_t$$

Introduction

- Differencing a cointegrated system may distort the determination of the number of factors due to the introduction of non-invertible moving average (MA) components and/or the trade-off introduced between the variances of the common and idiosyncratic components.
- Surprisingly, there has been little discussion in the literature on whether differencing in a univariate fashion affects the correct determination of the number of factors.
- Bai (2004) analyzes the performance of the Bai and Ng (2002) information criteria.

The model

- Consider the following DFM

$$y_{it} = p'_{i0} f_t + p'_{i1} f_{t-1} + \cdots + p'_{is} f_{t-s} + \varepsilon_{it},$$
$$y_{it} = p'_i(L) f_t + \varepsilon_{it},$$

where f_t is q dimensional and $p_i(L) = p_{i0} + p_{i1}L + \cdots + p_{is}L^s$.

The model

- We can write the previous model as static form

$$y_{it} = P_i F_t + \varepsilon_{it}$$

where

$$P_i = \begin{bmatrix} p_{i0} \\ p_{i1} \\ \vdots \\ p_{is} \end{bmatrix} \quad \text{and} \quad F_t = \begin{bmatrix} f_t \\ f_{t-1} \\ \vdots \\ f_{t-s} \end{bmatrix}.$$

- Note that the dimension of F_t is $r = q(s+1)$.

The model

- We assume that f_t is AR(h) process as follows

$$(I_q - B_1 L - \ldots B_h L^h) f_t = \epsilon_t,$$

or equivalently, F_t is a VAR with order depending on h and s. Let $\kappa = \max(h, s)$; then

$$\begin{pmatrix} f_t \\ f_{t-1} \\ \vdots \\ f_{t-\kappa} \end{pmatrix} = \begin{pmatrix} B_1 & B_2 & \cdots & B_{\kappa+1} \\ I_q & 0 & \cdots & 0 \\ 0 & I_q & \cdots & 0 \\ . & . & . & . \\ 0 & \cdots & I_q & 0 \end{pmatrix} \begin{pmatrix} f_{t-1} \\ f_{t-2} \\ \vdots \\ f_{t-\kappa-1} \end{pmatrix} + \begin{pmatrix} I_q \\ 0 \\ \vdots \\ 0 \end{pmatrix} \epsilon_t.$$

The model

- Note that $F_t^* = (f_t', f_{t-1}', \ldots, f_{t-\kappa}')$. We have

$$F_t^* = \Phi F_{t-1}^* + \eta_t,$$
$$\eta_t = R\epsilon_t.$$

- Consequently, when $s \geq h$, then $F_t \equiv F_t^*$ and when $h > s$ then, F_t is a $VAR(1, h-s)$, see Bai and Ng (2007).
- In this talk, we consider the case when F_t is a $VAR(1)$ process.

The model

- Then, the DFM is given by

$$Y_t = PF_t + \varepsilon_t,$$

$$F_t = \Phi F_{t-1} + \eta_t,$$

$$\varepsilon_t = \Gamma \varepsilon_{t-1} + a_t.$$

- $Y_t = (y_{1t}, \dots, y_{Nt})^{'}$ and $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{Nt})^{'}$ are $N \times 1$.
- $F_t = (F_{1t}, \dots, F_{rt})^{'}$ and $\eta_t = (\eta_{1t}, \dots, \eta_{rt})^{'}$ are $r \times 1$.
- η_t and a_t, are assumed to be Gaussian white noises with positive definite covariance matrices Σ_η and Σ_a respectively.
- $P = (p_1^{'}, \dots, p_N^{'})^{'}$ is $N \times r$ where, $p_i = (p_{i1}, \dots, p_{ir})$ is a $1 \times r$ vector.
- Φ and Γ are $r \times r$ and $N \times N$ matrices containing the autoregressive parameters.
- We assume $P^{'} P/N = I_r$ and $FF^{'}$ being diagonal.

Bai and Ng (2002) information criteria

- The most popular information criteria to select the number of factors in DFMs, proposed by Bai and Ng (2002), are based on a consistent PC estimator of P and F_t which is given by the solution to the following least squares problem

$$\min_{F_1,\ldots,F_T,P} V_r(P,F)$$

subject to $P^{'}P/N = I_r$ and $FF^{'}$ being diagonal,

with $V_r(P,F) = \frac{1}{NT}\sum_{t=1}^{T}(Y_t - PF_t)^{'}(Y_t - PF_t) = \frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N}\varepsilon_{it}^2 = \frac{1}{NT}tr(\varepsilon\varepsilon')$

Bai and Ng (2002) information criteria

- The solution to minimization problem is obtained by setting $\hat{P}$ equal to $\sqrt{N}$ times the eigenvectors corresponding to the r largest eigenvalues of YY' (denoted as Λ) where $Y = (Y_1, \ldots, Y_T)$. The corresponding PC estimator of F is

$$\hat{F} = (\hat{P}'\hat{P})^{-1}\hat{P}'Y = \frac{\hat{P}'Y}{N}$$

- Note that $YY' = \Lambda D \Lambda' \therefore \frac{\hat{P}'\hat{P}}{N} = \frac{\sqrt{N}\Lambda'\Lambda\sqrt{N}}{N} = I_r$ and $\hat{F}\hat{F}' = \frac{\sqrt{N}\Lambda' YY' \Lambda\sqrt{N}}{N} = D$

Bai and Ng (2002) information criteria

- PC factor extraction separates the common component, PF_t, from the idiosyncratic noises by averaging cross-sectionally Y_t such that when N and T tend simultaneously to infinity, the weighted averages of the ε_t converge to zero, remaining only the linear combinations of the factors. The asymptotic properties are given by Bai and Ng (2003).
- Stock and Watson (2011) point out that the factor space can be consistently estimated by PC even with certain types of breaks in the factors or time variation in the factor loadings. Intuitively, if under weak assumptions $\hat{F}$ consistently estimates a rotation of F, then, the factors can break or evolve in some limited fashion and the PC estimator will remain consistent.

Bai and Ng (2002) information criteria

- In order to determine r, Bai and Ng (2002) propose minimizing the following functions with respect to k, for $k = 0, \ldots, r_{\max}$,

$$IC_1(k) = \ln V_k(\hat{P}, \hat{F}) + k\frac{N+T}{NT} \ln \frac{NT}{N+T},$$

$$IC_2(k) = \ln V_k(\hat{P}, \hat{F}) + k\frac{N+T}{NT} \ln m,$$

$$IC_3(k) = \ln V_k(\hat{P}, \hat{F}) + k\frac{\ln m}{m},$$

where $m = \min \langle N, T \rangle$ and $r_{\max}$ is a bounded integer such that $r \leq r_{\max}$.

Differenced eigenvalues

- Onatski (2010) proposes an alternative procedure to select r, called edge distribution (ED), and shows that it performs well when the proportion of the variance attributed to the factors is small relative to the variance due to the idiosyncratic noises or when these are substantially correlated.
- This procedure is based on determining a sharp threshold, δ, which consistently separates the bounded and diverging eigenvalues of $\hat{\Sigma}_Y$.
- For any $j > r$, the differences $\hat{\lambda}_j - \hat{\lambda}_{j+1}$ converge to 0 while the difference $\hat{\lambda}_r - \hat{\lambda}_{r+1}$ diverges to infinity when both N and T tend to infinity.

Ratio of eigenvalues

- Ahn and Horenstein (2013) propose two further estimators of the number of factors based on the fact that the r largest eigenvalues of $\hat{\Sigma}_Y$ grow unbounded as N increases, while the other eigenvalues remain bounded. Their estimators are the value of k ($0 \leq k \leq r_{\max}$) that maximizes the following ratios

$$ER(k) = \frac{\hat{\lambda}_k}{\hat{\lambda}_{k+1}}$$

Ratio of eigenvalues

- and

$$GR(k) = \frac{\ln\left[V_{k-1}(\hat{P},\hat{F})/V_k(\hat{P},\hat{F})\right]}{\ln\left[V_k(\hat{P},\hat{F})/V_{k+1}(\hat{P},\hat{F})\right]} = \frac{\ln(1+\hat{\lambda}_k^*)}{\ln(1+\hat{\lambda}_{k+1}^*)},$$

where $\hat{\lambda}_0 = \frac{1}{m}\sum_{k=1}^{m}\hat{\lambda}_k/\ln(m)$ and $\hat{\lambda}_k^* = \hat{\lambda}_k/\sum_{j=k+1}^{m}\hat{\lambda}_j$.

Determining the number of factors after differencing

- Consider the DFM in which Φ and Γ are diagonal matrices which may have 1's in the main diagonal. Consequently, both the factors and the idiosyncratic noises can be either stationary or non-stationary random walks.

- After differencing the data in a univariate fashion, the DFM takes the following form (Bai and Ng, 2004)

$$\Delta Y_t = P\Delta F_t + \Delta\varepsilon_t,$$

$$\Delta F_t = (\Phi - I)F_{t-1} + \eta_t,$$

$$\Delta\varepsilon_t = (\Gamma - I)\varepsilon_{t-1} + a_t.$$

Determining the number of factors after differencing

- Denote by γ_i the i-th element in the main diagonal of Γ. If $|\gamma_i| < 1$, then the variance of the corresponding differenced idiosyncratic component is given by $\sigma^2_{e_i} = 2\sigma^2_{a_i}/(1+\gamma_i)$ where $\sigma^2_{a_i}$ is the variance of a_{it}. When $\gamma_i = 0.5$, the difference between the variances of ε_t and $\Delta\varepsilon_t$ is zero.

- However, if $\gamma_i < 0.5$, the variance of $\Delta\varepsilon_t$ is larger than that of ε_t while if $\gamma_i > 0.5$, it is smaller. The same relation can be established for the variances of the elements in F_t and ΔF_t with respect to ϕ_i, the i-th element in the main diagonal of Φ.

Determining the number of factors after differencing

- If ε_t is stationary, with autoregressive parameters smaller than 0.5 while F_t is non-stationary, then overdifferencing the idiosyncratic components may introduce distortions on the determination of the number of factors given that the relation between the variances of the common and idiosyncratic components is modified with the variances of ΔF_t being smaller and the variances of $\Delta\varepsilon_t$ being larger.
- IT IS THE COINTEGRATED CASE!

Conclusions

- Corona *et al.* (2017) show that both the variance and the dependence structure of the differenced idiosyncratic noises are important when determining the number of factors.
- If $r = 1$, the ER and GR procedures work well even in relatively small sizes and several structures of the idiosyncratic noises.
- The performance of all procedures deteriorates when $r = 2$. In this case, the ED procedure seems to work better.

Estimating non-stationary common factors

Introduction

- Once determined r, it is interesting to estimate the non-stationary common factors in order to known the common trends of the variables.
- In the context of cointegrated systems, Vahid and Engle (1993) Gonzalo and Granger (1995), among many others, provide procedures to estimate non-stationary common factors subject to cointegration results.
- In this talk, we focus on the context on large systems, studying the non-stationarity over the assumptions on the common factors in the DFM.

Introduction

- In the context of PC factor extraction, we consider the following methods of estimation, first, for data in levels:
    - Bai (2004) proposes factor extraction implementing PC to data in levels and derives the rates of convergence and limiting distributions of the estimated common trends and loading weights when the idiosyncratic components are stationary.
    - Corona *et al.* (2017) propose extend the procedure of Doz *et al.* (2011) in order to use the two step method to estimate non-stationary common factors when the idiosyncratic errors are $I(0)$.

Introduction

- Second, for first-differenced data:
  - Alternatively, when the idiosyncratic components are non-stationary, PC cannot be directly implemented to the original data as proposed by Bai (2004) and it is convenient to use the “differencing and recumulating” method proposed by Bai and Ng (2004).
- Barigozzi *et al.* (2016) project the original system onto the space spanned by the estimated loading as proposed by Bai and Ng (2004).

PC factor extraction

- When we use data in levels, the PC estimator of F is given by

$$\hat{F}^{PCL} = \frac{\hat{P}^{PCL'} Y}{N}$$

- Alternatively, when the common factors are $I(I)$, Bai (2004) proposes to use the restriction $FF'/T^2 = I_r$ with $P'P$ being diagonal.

PC factor extraction

- When the DFM is differentiated, the "differencing and recumulating" PC estimator is the following

$$\hat{F}_t^{PCD} = \sum_{s=2}^{t} \hat{f}_s, \quad \text{for} \quad t = 2, \ldots, T,$$

where $\hat{f}_t$ is the PC estimator obtained from ΔY_t.

PC factor extraction

- Bai and Ng (2004) and Barigozzi *et al.* (2016) show that $\hat{F}_t^{PCD}$ and/or $\hat{F}^{BLL}$ are a consistent estimator for a rotation of F_t up to a level shift regardless of whether the idiosyncratic component, ε_t, is $I(0)$ or $I(1)$.

Two-steps Kalman Smoother

- Considering the possibility of non-stationary common factors, consequently, when Φ may have 1's in the main diagonal, we propose to generalize the Doz et. al (2011) algorithm as follows

  1. Obtain the PC estimators of P and F_t with data in levels. Compute the idiosyncratic residuals $\hat{\varepsilon} = Y - \hat{P}^{PCL}\hat{F}^{PCL}$. Compute the covariance matrix of the idiosyncratic terms, $\hat{\Psi} = \text{diag}\left(\hat{\Sigma}_{\varepsilon}\right)$.

  2. Consider the j-th element of $\hat{F}_t^{PCL}$, $\hat{F}_{jt}^{PCL}$. For $j = 1$, we carry out the Augmented Dickey Fuller (ADF) test.

Two-steps Kalman Smoother

- ...

  3 If we reject the null hypothesis, we obtain by Ordinary Least Squares (OLS) the autoregressive coefficient, $\hat{\phi}_1$, the residuals as $u_{1t} = \hat{F}_{1t}^{PCL} - \hat{\phi}\hat{F}_{1t-1}^{PCL}$ and the variance of the factor disturbance as $\hat{\sigma}_{\eta_1}^2 = \sum_{t=1}^{T} u_{1t}^2 / T$. Estimate the initial state mean and variance as 0 and $\hat{\sigma}_{F_{10}}^2 = \hat{\sigma}_{\eta_1}^2 / (1 - \hat{\phi}_1^2)$, respectively.

  4 If we do not reject the null hypothesis, we fix the autoregressive parameter $\hat{\phi}_1 = 1$. Compute the residuals as $u_{1t} = \Delta\hat{F}_{1t}^{PCL}$. Calculate the variance of the factor residuals, $\hat{\sigma}_{\eta_1}^2 = \sum_{t=2}^{T} \Delta\hat{F}_{1t}^{PCL^2} / (T-1)$. Use a diffuse initial state mean and variance as 0 and $\hat{\sigma}_{F_{10}}^2 = \kappa$ (for instance $\kappa = 10^7$).

Two-steps Kalman Smoother

- ...

  5 Repeat steps 2 to 4 for $j = 2, \ldots r$, to obtain $\hat{\Phi} = \text{diag}(\hat{\phi}_1, \ldots, \hat{\phi}_r)$, $\hat{\Sigma}_\eta = \text{diag}(\hat{\sigma}^2_{\eta_1}, \ldots, \hat{\sigma}^2_{\eta_r})$ and $\hat{\Sigma}_{F_0} = \text{diag}(\hat{\sigma}^2_{F_{10}}, \ldots, \hat{\sigma}^2_{F_{r0}})$.

  6 Use $\hat{P}$, $\hat{\Psi}$, $\hat{\Phi}$, $\hat{\Sigma}_\eta$ and $\hat{\Sigma}_{F_0}$ in the KS to obtain the estimators of the common factors $\hat{F}^{KSL}$.

Two-steps Kalman Smoother

- Furthermore, we can apply KS procedure to ΔY_t and obtaining the "differencing and recumulating" KS estimator, denoted as $\hat{F}^{KSD}$.
- The asymptotic properties of $\hat{F}^{KSD}$ are known following to Doz *et al.* (2011) even if, the idiosyncratic errors are $I(1)$.
- Barigozzi and Luciani (2017) propose a similar approach using EM algorithm, however, they do not consider the "differencing and recumulating" KS estimator.

Conclusions

- In the context of non-stationary DFMs, when the idiosyncratic component is $I(0)$, the procedures that extract the common factors using data in levels perform better that the procedures that use first-differenced data.

- Corona *et al.* (2017) show that, when the idiosyncratic errors are non-stationary, the approaches based on estimating the common factors using non-stationary time series in levels do not perform well and that the procedures based on first differences should be used.

Dynamic factor structure of team performances in Liga MX

Dynamic factor structure of team performances in Liga MX

- Corona, González-Farías and Muriel (2018) analyze the team performance of the Mexican Football League (Liga MX), measured as the percentage of the total points obtained during each short tournament using DFMs.
- The estimation of the common components is carried out with Principal Components (PC) and the stochastic nature of the DFM is studied through Panel Analysis of Non-stationarity in Idiosyncratic and Common Components (PANIC).

Dynamic factor structure of team performances in Liga MX

- We define the observed variables, $\{y_t = (y_{1t}, \ldots y_{Nt})'\}$, as the percentage of points obtained for each team (indexed by $i = 1, \ldots, N$) and tournament (indexed by $t = 1, \ldots, T$). We use data starting spanning from 1996 to 2017 and including two observations per year. Since 1996, the Mexican league follows a schedule of two tournaments a year.

Dynamic factor structure of team performances in Liga MX

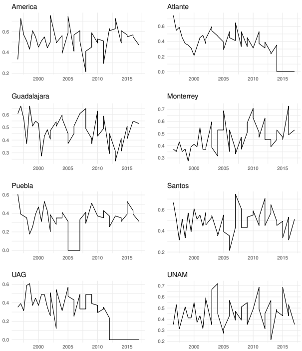


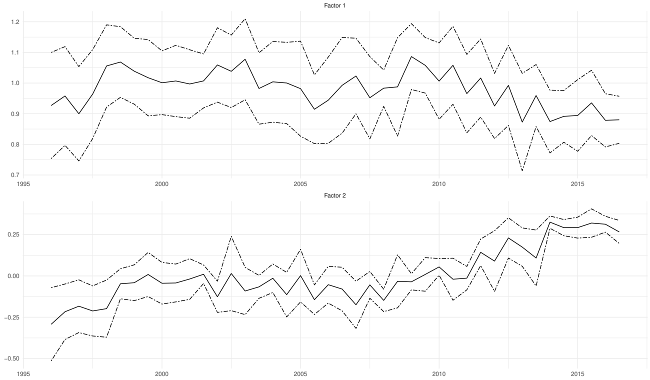
Figure: Percentage of the total points obtained by each of the teams included in the study for the tournaments starting in Invierno 1996 and ending in Clausura 2017

Dynamic factor structure of team performances in Liga MX

- We observe that $\widehat{r}_{IC_1}$, $\widehat{r}_{IC_2}$ and $\widehat{r}_{ED}$ detect 2 common factors while the ratios of adjacent eigenvalues determine $\widehat{r} = 1$. On the other hand, $\widehat{r}_{IC_3} = r_{\max}$. Taking the analysis of finite sample performance of Corona *et al.* (2017) into account, we conclude that $\widehat{r} = 2$.
- PANIC concludes that the idiosyncratic component is $I(0)$ and the common factors are or both $I(0)$ (trend specification) or only $F_{2t} \sim I(1)$ (constant specification).

Dynamic factor structure of team performances in Liga MX

Figure: Estimated common factors (solid lines) and their 95% confidence intervals (dashed lines)

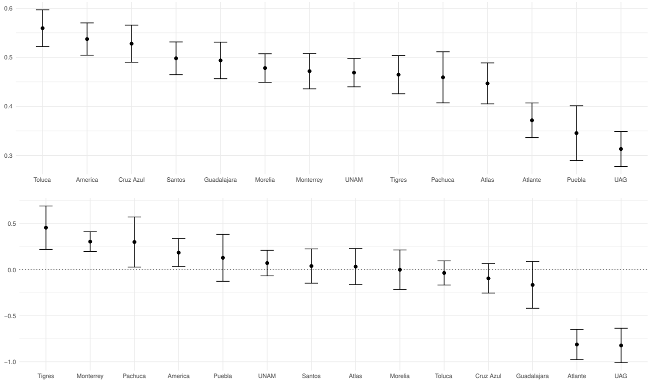


Dynamic factor structure of team performances in Liga MX

- The behavior of the common factors is intresting: The first common factor has a constant mean around zero, while the second factor exhibits a slightly positive trend, mainly from 2011 onwards. Notice that the confidence intervals are tight reflecting a rather small estimation uncertainty.
- The contribution of each team to the factors is depicted in the following Figure, where confidence intervals estimated according to Bai (2003), are also given.

Dynamic factor structure of team performances in Liga MX

Figure: Loading weights and their 95% confidence intervals. The confidence intervals were calculated following Bai (2003).

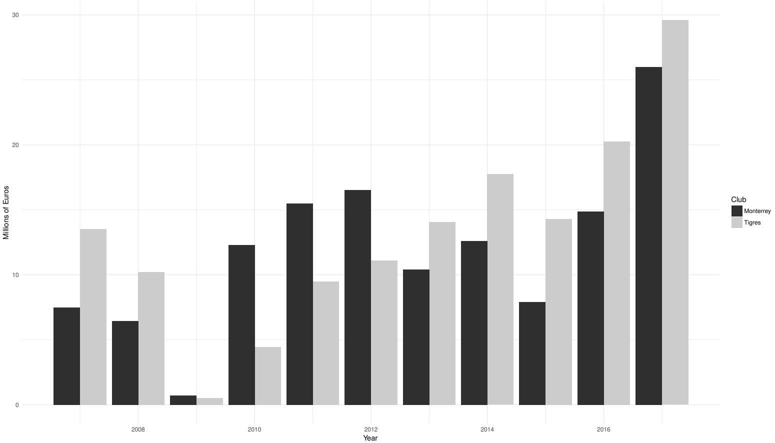


Dynamic factor structure of team performances in Liga MX

- We observe that the first common factor is dominated by all teams, however, the teams with the most contribution are Toluca, America, Cruz Azul, Guadalajara and Santos, i.e., the teams with the most points during “short tournaments”.
- The second common factor is interesting given that it is positively dominated by Tigres and Monterrey, the emergent teams since 2009. Currently, discussion is taking place over whether these two teams are the “new bigger” teams, competing with traditional teams such as America, Guadalajara, Cruz Azul and UNAM.

Dynamic factor structure of team performances in Liga MX

Figure: Millions of euros more invested in average than other teams. Light color is Tigres and dark color Monterrey.



Dynamic factor structure of team performances in Liga MX

- The results indicate that we determine two common factors, the first dominated by the traditional teams and the second, mainly, by emergent teams as Tigres and Monterrey and also, by relegated teams as Atlante and UAG.
- It is interesting to the second common factor is non-stationary when we use a constant in PANIC, which is influenced by the performance of these latest teams. On the other hand, the idiosyncratic component is stationary, implying that the individual behavior of each teams has constant mean and variance.

Dynamic factor structure of team performances in Liga MX

- Consequently, it is complicated to establish whether the DFM is cointegrated or not. If we assume cointegrated (second common factor is non-stationary), the non-stationarity performance of each team can be predicted by the performance of another team by the teams associated with significant factor loadings, as Tigres, Monterrey, Pachuca and America, then share a common trend.

A Dynamic Factor Model for the Mexican Economy: Are Common Trends Useful When Predicting Economic Activity?

Forecasting the Mexican economic activity

- Corona, Gónzalez-Farías and Orraca (2017) propose to use the common trends of the Mexican economy in order to predict economic activity one and two steps ahead. We use 211 macroeconomic and financial variables $I(1)$ from March 2005 to April 2016.
- We exploit the cointegration properties of the macroeconomic time series, such that, when the series are $I(1)$ and cointegrated, there is a factor representation, where the common factors are the common trends of the macroeconomic variables.

Forecasting the Mexican economic activity

- We apply the criteria described in this talk to detect the number of factors. We use an $r_{\max} = 11$. The results indicate that $\hat{r}_{ER} = \hat{r}_{GR} = 2$ and $\hat{r}_{ED} = 5$. Given that Onatksi (2010) is more robust in presence of non-stationarity, we work with this number of factors.

Forecasting the Mexican economic activity

- In order the predict the Mexican economic activity, we use a Factor Augmented VAR (FAVAR) representation as follows:

$$\begin{pmatrix} x_t \\ F_t \end{pmatrix} = \Phi(L) \begin{pmatrix} x_{t-1} \\ F_{t-1} \end{pmatrix} + v_t,$$

where

$$\Phi(L) = \begin{pmatrix} \Phi_{11}^{(k)} & \Phi_{12}^{(k)} & \Phi_{13}^{(k)} & \dots & \Phi_{1r+1}^{(k)} \\ 0 & \Phi_{22}^{(k)} & 0 & \dots & 0 \\ 0 & 0 & \Phi_{33}^{(k)} & & 0 \\ \vdots & \vdots & & \ddots & \\ 0 & 0 & 0 & & \Phi_{r+1r+1}^{(k)} \end{pmatrix} L^{k+1}.$$

Intuitively, common economic trends that summarize its long-run behavior are used to predict a specific macroeconomic variable.

Forecasting the Mexican economic activity

- We verify that the common trends of a large dataset of the economic variables are cointegrated with economic activity using PANIC and cointegration analysis.

- With the aim of selecting the forecast model, we consider all possibilities of FAVAR models, $\sum_{i=1}^{r} {}^{r}C_i \times 396$, where ${}^{r}C_i$ is the binomial coefficient $\binom{r}{i}$ and 396 is obtained as the product of 11 seasonal dummies (3 to 12 and none), 3 deterministic specifications in the FAVAR model (none, constant and constant-trend) and 12 lags (1 to 12).

Forecasting the Mexican economic activity

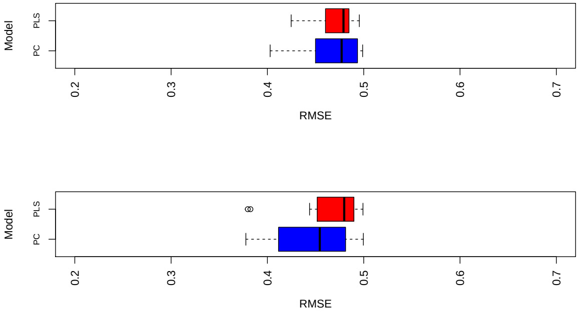


Figure: Box plots for the RMSE of out-sample forecasts for the selected models. The top panel is $h = 1$ and the bottom panel $h = 2$. The blue color refers to PC and the red color to PLS.

Forecasting the Mexican economic activity

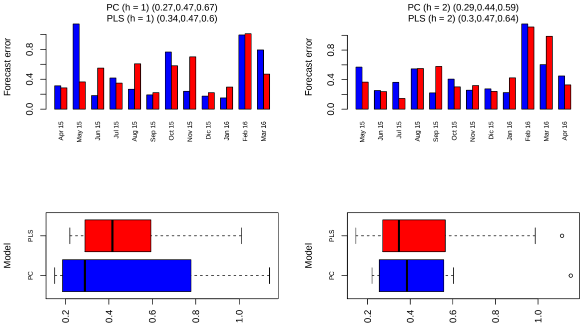


Figure: Forecast Errors for FAVAR models: April 2015 - 2016. PC (red color) and PLS (blue color). The top left panel plots $h = 1$ and the bottom right panel plots $h = 2$.

Forecasting the Mexican economic activity

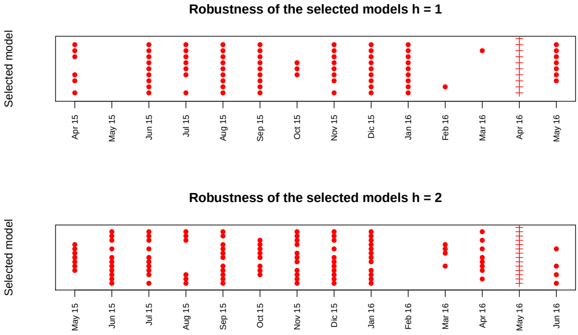


Figure: Robustness of selected models. Red points indicate that the model satisfies the empirical threshold in the specific forecast month. A red cross indicates that there is not information to compute the statistic.

Forecasting the Mexican economic activity

- We find that 211 macroeconomic and financial variables can be summarized in at most $\hat{r} = 5$ common trends, which are cointegrated with Mexican economic activity. Furthermore, we use the common trends in a FAVAR model with the aim of predicting Mexican economic activity.

- An important conclusion is that the macroeconomic common trends can be used in more sophisticated models in order to reduce the forecast errors. Additionally, note that in this study we used data in levels in order to estimate the common factors.

Acknowledgements

The author acknowledges the partial financial support from CONACYT-CB-252996.