Outline

- Motivation
    - PLS
    - PLSAR
- Quick overview of VAR Models
- Definition for VAR-PLS
- Boostrap for VAR-PLS
- Conclusions

Motivation

- PLS is a techique that has bee proven its impact on many applications such as quality control starting with the *Chemistry*, batch proceses, medical images analysis, microarrays, path modeling, classification, discrimination, spacio-temporal PLS models just to mention some, with authors such as McGregor, Nomikos, MacIntosh, V. Esposito Vinz, P. Garthwaite, and so on

- The method can be used in univariate and multivariate data as well

- *It has been shown that gives better prediction even when the stantard assumptions are met*

- Phillip Hans Franses (2006) propose a methodology to construct the forcast h steps ahead in an optimal way, through an autorregresive order p: "An Autorregresive Parcial Least Square (PLS)" denote as $PLSAR(h, p)$

Our case of interest

- Develop a model to predict the Mexican inflation, as precise as possible.

- The model has to consider, as the principal source of the mexican inflation, the grow and the variation on the monetary condition of the country

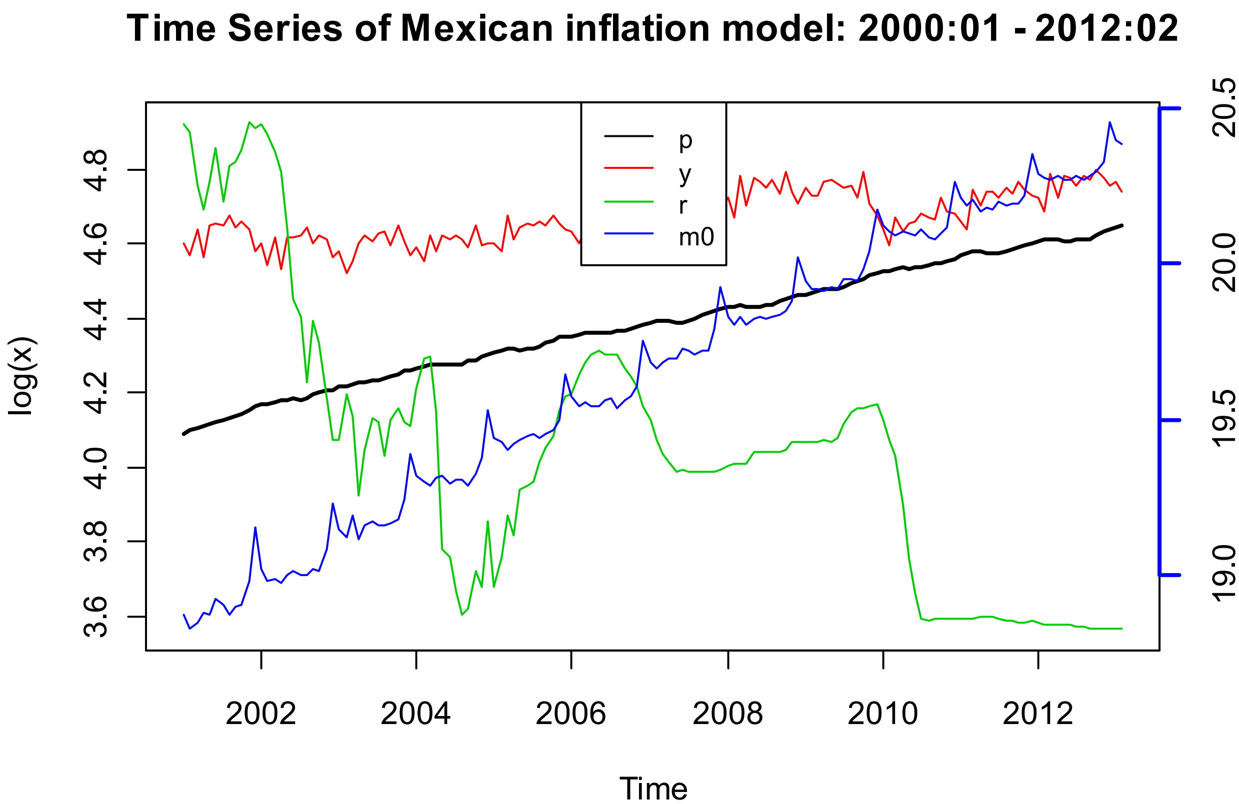
- Irrespective of all possible discussions, there seems to be a common undersanting into belive that the inflationary process, in the lung run, is a purely monetary fenomena

- Here we are not taking the disussion on the existence or not of such relationship but we will show its empirical properties with a model that is tested out of the sample via its error prediction measure.

We work with 4 indexes (we bulit those) from January 2000 to February 2012:

- **p**: Consumer price index
- **m0**: Monterary base
- **r**: Equilibrium interest rate (28 días)
- **y**: Industrial production index

Relacionar empíricamente la variable de precios, que a su vez es una función de la tasa de inflación (inflación mensual, inflación interanual, acumulada, etc.) con el resto de las variables permitiendo las relaciones multivariadas existentes, generando así un $VAR(p)$ de rango completo y/o un $VECM$ para el caso cointegrado.



Podemos apreciar que la serie de precios lleva una clara tendencia creciente en niveles, el índice monetario presenta una estacionalidad característica en todo el periodo de tiempo, la producción económica con tendencia de menor pronunciación que la serie de precios la cual que a partir del 2010 exhibe cierta recuperación respecto a los niveles observados en la primera mitad del gráfico. La tasa de interés claramente ha tenido un periodo de estabilidad a partir del segundo semestre del 2010.

- We generalized the work propossed by Franses in the following way:
  (here we just show some comparative results)

  1. Give a multivariate representation based on the flexibility of the VAR models, model that we will call $VAR-PLS(h,p)$
  2. Extend the model to consider deterministic variables (dummies, trend, etc.) as well as exogenous variables
  3. Boostrap predicction intervals
  4. Compare the forecast capabilities between a $VAR-PLS(h,p)$ and a forecast VAR model explicity buit for (integral predictor method)

$PLSAR(h, p)$

Franses plantea la comparación entre tres formas de hacer pronósticos bajo un AR(p).

- A single model for all horizons , an interative procedure will come on hand (escribir el modelo sin los gorros y con el termino de error y con rho no p)

$$AR(p): \; y_{T+h} = \hat{\mu} + \hat{p}_1 y_{T+h-1} + \hat{p}_2 y_{T+h-2} + \cdots + \hat{p}_p y_{T+h-p} + \hat{\varepsilon}_t$$

❖ El modelo $AR(p)$ es la forma clásica de realizar los h pasos hacia adelante cuyos parámetros son estimados generalmente por Mínimos Cuadrados Ordinarios (OLS)

- One model for each horizon, the variance will vary within each horizon (escribir modelo sin los gorros y con el término de error y con rho no p)

$$AR_h(p): \; y_{t+h} = \hat{\mu} + \hat{p}_{1,h} y_t + \hat{p}_{2,h} y_{t-1} + \cdots + \hat{p}_p y_{t-p} + \hat{\varepsilon}_{h,t}$$

The $AR_h(p)$ is an alternative to the $AR(p)$ because

❖ OLS minimize the sum of square of $\hat{\varepsilon}_t$ but there is no way to assure it will remain minimum for all the h steps in the future

  ▪ One can count with different models for each step

❖ For stationaty time series recall the the forecast of an $AR(p)$ model quickly converge to the inconditional mean (and variance, for the interval predicton error) clearly depending on $h \geq p$

Note: For more details on this type of models see : Pesaran & Pick (2010), Marcellino, Stock & Watson (2004), Carreiro, Kapetorios & Marcellino (2010), Tiao & Xu (1993) among others.

- Something in between: PLSAR, este modelo se encuentra situado entre un $AR(p)$ que pronostica todos los h pasos adelante y diferentes modelos AR para cada horizonte, formuló un modelo para h pronósticos denominado

$$PLSAR(h, p): Y = XB_{PLS} + U$$

❖ It is clear that exists adjacent correlation between the time series, and neither one of the above models take them into account. In other words, we know that y_t and y_{t-j} are correlated and so are y_{T+h} and y_{T+h-j}, . Therefore we would like to jointly predict $(y_{t+h}, y_{t+h-1}, y_{t+h-2}, \dots, y_{t+1})$ through $(y_t, y_{t-1}, y_{t-2}, \dots, y_{t-p})$. PLS is a techinique very atractive to do so.

❖ Franses propose to arrange the information as

- $(y_t, y_{t-1}, y_{t-2}, \dots, y_{t-p})$ as the predictor X matrix

and

- $(y_{t+h}, y_{t+h-1}, y_{t+h-2}, \dots, y_{t+1})$ as the predicted Y matrix

- Applied the PLS algorthim, to get the latent variables with the relevant information given in X and Y.

His simulations shows that the $PLSAR(h,p)$: is quite comptetive with respect to the classical models in the literature.

PLS

- PLS can be tract from different stand points, to us the relationship between its linear expresssion will be the best one, in order to related with a Vector Autorregresive model
$$Y = XB + U$$

donde Y es Nxk, X es una matriz nxN, B es una matriz de $(N+1)xk$ y U es una matriz de Nxk.

The basic procedure maximize the $cov(X\alpha, Y\beta)^2$ $under$ certain restrictions

$$\alpha'(S^*{}_{xx} + \lambda_x)\alpha = 1 \text{ \& } \beta'\left(S^*{}_{yy} + \lambda_y\right)\beta = 1$$

where

$$S^*{}_{xx} = (1-\lambda_x)S_{xx} \text{ \& } S^*{}_{yy} = \left(1-\lambda_y\right)S_{yy}$$

- $(X\alpha, Y\beta)$ are linear combinations of the variables that maximize the covariance or actually the squere covariance (the sign is not important just the direction)
- S_{xx} y S_{yy} the varince covariance matrices; $\beta'\beta = 1$ & $\alpha' S_{xx}\alpha = 1$

we maximze the the objective function:

$$\mathcal{L} = \left(\alpha' S_{xy}\beta\right)^2 - \gamma\{\alpha'(S^*{}_{xx} + \lambda_x)\alpha - 1\} - \mu\left\{\beta'\left(S^*{}_{yy} + \lambda_y\right)\beta - 1\right\}$$

After some algebra we get the *scores* for X y Y, $t = Xw = Ew$ and $u = Yq = Fq$

Normalizing the scores $t = t/\sqrt{t't}$ and after simplifications and more algebra we get the *loadings* for X and Y: $p = E't$ and $q = F't$. Writing in matrix form w, t, p y q we get $R = W(P'W)^{-1}$ that basically recovers the basic information for the estimated parameters

$Y = XB + U$ so $Y(gorro) = XB_{PLS}$ where $B_{PLS} = R(T'T)^{-1}T'Y = RQ'$

Note 1: For a nice introduction see : P.H. Garthwaite (1994). An Interpretation of Partial Least Squares. JASA Vol 89, No 425, pp 122-127 and Agnar Hoskuldsson (1988). PLS Regression Methods. Journal of Chemometrics, Vol 2, pp 221-228.

Note 2: Franses shows that if the $\mathrm{B}_{\mathrm{PLS}}$ matrix has full rank it implies a different model for each of the columns of Y , and hence a model like (2). In the exceptional case that $\mathrm{B}_{\mathrm{PLS}}$ has rank 1, then the AR(p) appers.

Vectores Autorregresivos y Mínimos Cuadrados Parciales

A $VAR(p)$ processes is defined as

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + CD_t + u_t \tag{4}$$

A_i: Coefficient Matrices for $i = 1,2,\dots,p$
ε_t: White noise process with covariance matrix $E[u_t u_t{}'] = \Sigma_u$
C: Possible deterministic regressor coefficient matrix
D_t: Appropied vector for deterministic variables

It is well know that a $VAR(p)$ model can be written as a $VAR(1)$ in the following way

$$Y_t = AY_{t-1} + V_t \tag{5}$$

$$Y_t = \begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{bmatrix}$$

$$A = \begin{bmatrix} A_1 & A_2 & \cdots & A_{p-1} & A_p \\ I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{bmatrix}$$

$$V_t = \begin{bmatrix} u_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

- If the eigenvalues of A are less than one, then the $VAR(p)$ is stable

---

- We use the VAR representation to determine the order of the model

The general procedure is quite standard. Order $p = 0,\dots,p_{max}$ and choose the value of p than minimize some criteria. The criteria are usually written as:

$$IC(p) = ln\left|\hat{\Sigma}(p)\right| + C_T.\varphi(K,p)$$

where $\hat{\Sigma}(p) = T^{-1}\sum_{i=1}^{T}\hat{u}_t{}'\hat{u}_t$, C_T indexed sequence of the size of the series, T and $\varphi(K,p)$ is a penalty function that involves the order of the $VAR(p)$.
The most common information criterias are: Akaike (AIC), Schwarz-Bayesiano (BIC), Hannan-Quinn (HQ) and Final predicction error (FPE):

$$AIC(p) = \left|\hat{\Sigma}(p)\right| + \frac{2}{t}pK^2$$
$$BIC(p) = \left|\hat{\Sigma}(p)\right| + \frac{lnT}{t}pK^2$$
$$HQ(p) = \left|\hat{\Sigma}(p)\right| + \frac{2lnlnT}{t}pK^2$$
$$FPE(p) = \left(\frac{T+p^*}{T-p^*}\right)^K det\left(\hat{\Sigma}(p)\right)$$

- The AIC asymptotically overestimate the order of the model with a positive probability whereas BIC and HQ are consisten estimator of the order if, the real value of is less than or equal than p_{max}

Notemos que al igual que el caso univariado, podemos con (4) pronosticar recursivamente de la siguiente manera:

$$y_{T+h\backslash T} = A_1 y_{T+h-1} + \cdots + A_p y_{T+h-p} + CD_{T+h}$$

- The estimation of the A_i matrices can be done by OLS

$$vec(\hat{A}) = \begin{bmatrix} \hat{A}_1 \\ \vdots \\ \hat{A}_p \end{bmatrix}$$

- Under statiobarty and ergodicity condition for the VAR models (see Hamilton (1994), Lutkepohl (1991) among otheres), $vec(\hat{A})$ a consistent and asymtotically distributued with covariance matriz given by

$$\widehat{avar}\left(vec(\hat{A})\right) = \hat{\Sigma} \otimes (Z'Z)^{-1}$$

where

- $\hat{\Sigma} = \frac{\sum_{i=1}^{T}\hat{\varepsilon}_t{}'\hat{\varepsilon}_t}{T-K}$
- $\hat{\varepsilon}_t = Y_t - \hat{A}'Z_t = Y_t - \hat{A}'Y_{t-1}$ the residual at time t
- The $i-th$ element of $vec(\hat{A})$ is asyntotically normal (for an a stable VAR) and the stanrdad error are the square roots of the diagonal elementes of $\hat{\Sigma} \otimes (Z'Z)^{-1}$.
- The t -test are asymtotically valids for the estimated coeficients

---

- Other important situation for the VAR models is the presencia of one or more unit roots y_j.
- From the point of view of the economic theory it mean th study of a lung run behavior plus the temporal dynamic of the series.

- **Cointergration:** The componetes of y_t are cointegrated of order d, b, denoted by $y_t \sim CI(d,b)$ if
  a) All the componentes of $y_t \sim$ are $I(d)$; and
  b) There exisit a vector $\beta(\neq 0)$ such that $z_t = \beta' y_t \sim I(d-b), b > 0$. The vector β is called the cointegrated vector

- **Error correction:** The bivariate vector $y_t = (y_{1t}, y_{2t})'$ has cointegrated vector $\beta = (1, -\beta_2)'$, and $\beta' y_t = y_{1t} - \beta_2 y_{2t} \sim I(0)$ then, exist an error correction representation

  a) $\Delta y_{1t} = \alpha_1 + \gamma_1 (y_{1t-1} - \beta_2 y_{2t-1}) + \sum_{i=1}^{K} \psi_{1,i} \Delta y_{1t-i} + \sum_{i=1}^{K} \psi_{2,i} \Delta y_{2t-i} + u_{1t}$

  b) $\Delta y_{2t} = \alpha_2 + \gamma_2 (y_{1t-1} - \beta_2 y_{2t-1}) + \sum_{i=1}^{L} \xi_{1,i} \Delta y_{1t-i} + \sum_{i=1}^{L} \xi_{2,i} \Delta y_{2t-i} + u_{2t}$

The $VAR(p)$ model can be written

$$\Delta y_t = \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + \ldots + \Gamma_{p-1} \Delta y_{t-p+1} + CD_t + u_t$$

$$\Gamma_i = -(A_{i+1} + \cdots + A_p), \quad i = 1, \ldots, p-1$$
$$\Pi = -(I - A_1 - A_2 - \ldots - A_p)$$

A Vector correction model (VECM) :transitorio

Or as:

$$\Delta y_t = \Pi y_{t-p} + \Gamma_1 \Delta y_{t-1} + \ldots + \Gamma_{p-1} \Delta y_{t-p+1} + CD_t + u_t$$

$$\Gamma_i = -(I - A_1 - A_2 \ldots - A_i), \quad i = 1, \ldots, p-1$$
$$\Pi = -(I - A_1 - A_2 - \ldots - A_p)$$

VECM : Lung run

The matrix Π represent the lung run behavior and if:

1. $rk(\Pi) = n$, the n linear combinations are stationary; in other word the VECM is no more tha VAR model in levels

2. $rk(\Pi) = 0$, there is no linear combination that makes Πy_{t-1} stationary, except for the trivial solution i.e., it becomes and $VAR(p-1)$ in first differences

3. 0< $rk(\Pi) < n$, the most interesting case, we can write $\Pi = \alpha\beta'$ (nxr) and $\beta' y_{t-1}$ is stationary. Each column of the matrix β represent a lung run relationaship

- **If the objective is to forecast series that are integrated or cointegrated doing with through a VAR representation is the most appropiated** (see Lutkepohl 2006)

From the application:

- In the infaltion example we specify the orden of the model through the final error prediction criteria, it was $p = 3$.
- We also performe Johansen test to determine the presence of lung run relationship. Finding the following relathionship which it was significant to 1%, level :

$$\boldsymbol{p + 9.68 - 0.43m0 - 0.89y + 0.10r = 0}$$

- This relationship is congruent with the economic theory behind it. The inflationary movemente increases for the monetary grow, the exced on demand and the reduction on the money cost.
  traspaso inflacionario está impulsado por el crecimiento monetario, exceso de demanda y la reducción del costo del dinero.

VAR-PLS

1. The VAR model will the the DGP
2. The VAR model will provide the autoregressive process that we will use to built the PLS regression
3. We built then the matrices in at natural way as:

- For the X matrix we include the lag vector: $X = Y_{t-1} = \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p} \end{bmatrix}'$ and
- For the Y the observed until time t $Y = Y_t = \begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{bmatrix}'$ using the X with all the lags consider for the DGP.

- We can introduce exogenous variables with a C matriz, then $X = Y_{t-1}^{*} = \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p} \\ D_t \end{bmatrix}$ and the coeficient matrix with look like $A^{*} = \begin{bmatrix} A_1 & A_2 & \cdots & A_{p-1} & A_p & C \\ I & 0 & \cdots & 0 & 0 & 0 \\ 0 & I & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 & 0 \end{bmatrix}$

- Those are the basic ingredients for a VAR-PLSX that will allows to predict *h* steps ahead

For the $VAR(p = 3) - PLS(h = 24, j)$

- We kept 24 observation to have a long horizon of possible comparisons
- We consider dummies for the monthly effects
- For the optimal p óptimo we estimate VAR-PLS, we use R to fit Yt (119X 4) = (Yt-1,D) (119x 23) * B(23X4) + U (23X4) and predict for the $VAR(3) - PLS(24, j)$ in a recursive way as in the VAR(p). According with Franses exercise, the last component agrees with the VAR(p) OLS estimate
- For the $pK + g = 23$ components and the 24 out of the sample steps we get the $MAPE$ to make the comparison

Para el modelo $VAR(p = 3)$,

- Combining all the variables estimate our $VAR_j(p), j = 1, \ldots, 3696$

- For the 24 steps out of the sample we use 7 different criterias to meausre the behavior of the forecast (Hyndman & Koehler 2006):

  - MAPE: Mean Absolute Percentage Error
  - MdAPE: Median Absolute Percentage Error
  - RMSPE: Root Mean Square Percentage Error
  - RMdSPE: Root Median Square Percentage Error
  - MRAE: Mean Relative Absolute Error
  - MdRAE: Median Relative Absolute Error
  - GMRAE: Geometric Mean Absolute Error

  Note: For the last 3 we need to work witha benchmark model (autoregresie order 1)) and for I i =1,.., h (h = 24) obtain the statistic:

- $test = \frac{Y_{t+i} - Y^f_{t+i,VARj(p)}}{Y_{t+i} - Y^f_{t+i,AR(1)}}$

- With the 7 criteria, integrate to one by taking the 0.5 quantil for the horizon consider, repite it for their upper and lower confidence intervals.

Intervalo de predicción: VAR-PLS

- We also have to construct the prediction intervals for the VAR-PLS model.

We use a similar procedure to the one proposed by Pascual, Ruíz and Fresoli (2011). *Bootstrap forecast of multivariate VAR models without using the backward representation*. Working Paper 11-34, Statistics and Econometrics Series. They use the seminal ideas of Kim (2001and some results from a previous work, Pascual, L., J. Romo, and E. Ruiz (2004a). Bootstrap predictive inference for ARIMA processes, Journal of Time Series Analysis, 25, 449-465

- For the VAR model they proposed a method that coupes with:

    1) The uncertanty give by the estimation of the parametrer, building confidence regions using their bootstrap method

        a) *This regions are valid under Gaussian assumptions (Lütkepolh et al, 1991), even though do not reflect, for small sample size, the asymmetric distribution of the predictd values (under estimated parameters)*

    2) The backward representation makes calculations quite complicate more in the case of the VAR(p) representation and p taking vaues greater than 5 for example, which is very common.

        a) *"Pascual et al shows that the backward representation can be avoided without loosing the good properties of the boostrap procedure"*

- Of course we needed to adequate the procedure to the procedure to the VAR-PLS representation

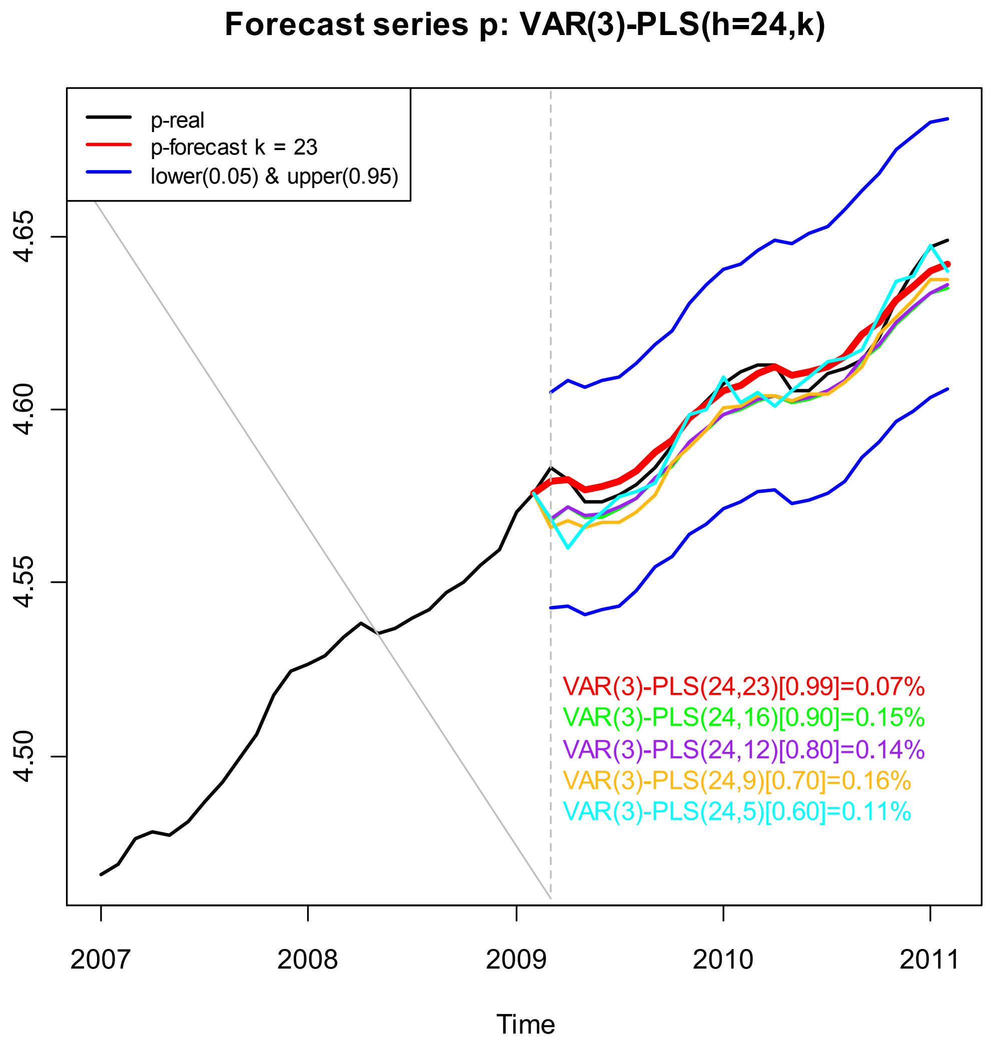
    Step 1. Fit the model to get $Y_t = X_t\hat{B}_{PLS}$
    Step 2. Obtain the standarized residuals $\hat{U}_t^*$ and get the emprical distribution of the residuals,
    Step 3. With the p initial values $Y_0 = \{Y_p, \dots, Y_1\}$ and the values obtain in Step 1 and 2 generate Y_t^*, the boostrap values, through were the $\hat{U}_t^*$ are independent drwas from its empirical distribution :

$$Y_t^* = X_t\hat{B}_{PLS} + \hat{U}_t^* \quad t = 1, \dots, n-p$$

    Step 4. We proceed in this way to get $\hat{Y}_{T+h}^*$ replicating steps 2 to 4 for $n = 1, \dots, N$
    Step 5. For each one of the nth variables and the set of N forecast we get:

$$CI_{T+h} = \{y_{n,T+k} | y_{n,T+k} \in [q_B^*(\tau), q_B^*(\tau-1)]\}$$

where $q_B^*(\tau)$ is the τth percentil of $G_{n,B}^*(x) = \#\left(y_{n,T+k}^{*(b)} \leq x\right)/N$



- From the economic point of view, the approximation is excelent.
- We observe tahat either using, $k = 23$ (0.99 of the variability) with an error precentage of 0.07% or with 70% explanation with a error percentage of 0.16%, the real value and the predicted one, for practical purposes are almost identical.
- The boostrap interval is very well behave.

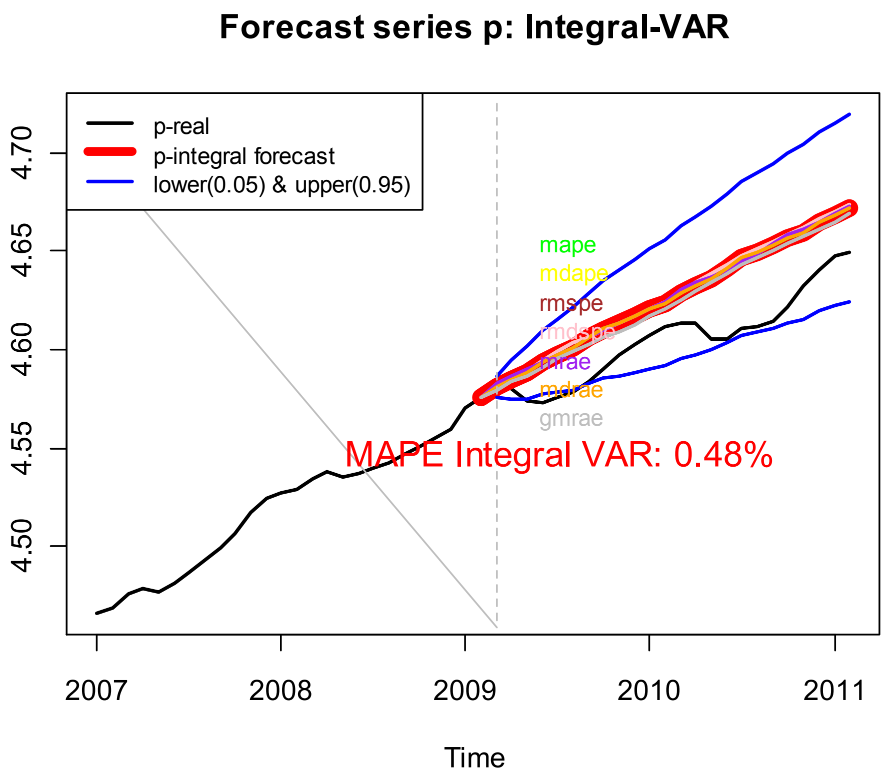
NOTA: The objective is to forecast the price, however since it ia a multivariate modelo we also get forecast for the other 3 variables with forecast error(Average of MAPE) of : **0.20% for the**

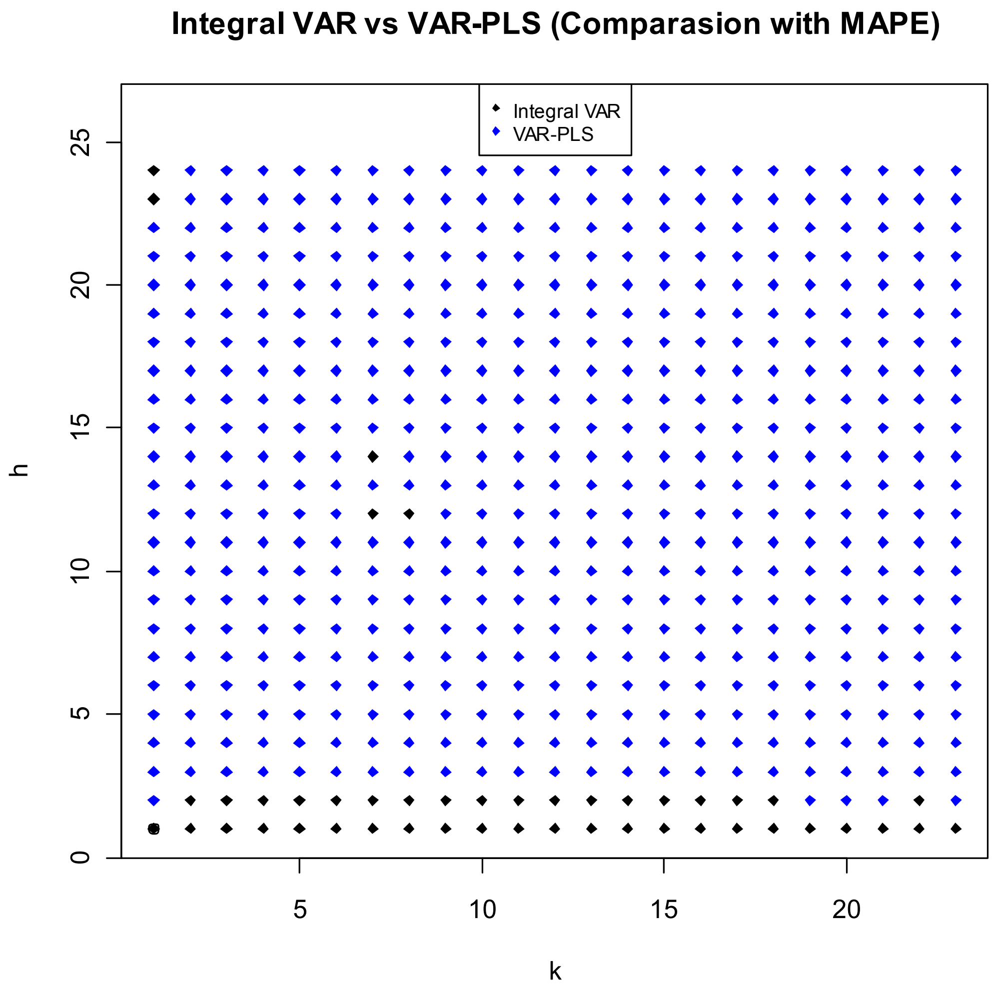
monetary base, 0.60% for industry production index and 12.03% for the equilibrium interest rate.

____

For the integral-VAR, the optimum $VAR_j(p)$ are:

Criteria	**MAPE**	**MdAPE**	**RMSPE**	**RMdSPE**	**MRAE**	**MdRAE**	**GMRAE**
Statistic	**0.16**	**0.12**	**0.19**	**0.12**	**0.16**	**0.12**	**0.11**
Variable	**r**	**r**	**r**	**r**	**r**	**r**	**r**
lags	3	2	3	2	2	3	5
Stacionality	9	11	9	11	6	9	11
Specification	none	none	none	none	none	none	none





In average 91.67% **the PLS representation of the VAR over all the components. It make sense that the last** components are less effective.

The VAR-PLS seems to be an atractive competitor of the, **integral VAR which is constructed for prediction purposes, with the advantage of a confidence intervals that include the uncertainty due to the parameter estimation.**

Conclusiones

- We present an alternative forecast alghoritm for a multivariate framework, that takes into account the dependecies among the series (y_{t+j} with y_t), trhough a PLS model written as a linear model with X having the VARX representation.

- The empiracl result for the mexican inflation are quite appropiate form an economical point of view
  - VAR-PLS was built with p given from the fit of a VARX, estimating in this case, all the pK possible components

- We construct the prediction interval using a Bootstrap method

- The VAR-PLS is an atractive multivariate technique to forecast this type of data.

Future work:

- Look for the implicit relationship between cointegration (CCA) and PLS
- Build a *PLS-VAR* without fitting the VAR model and compare its forcast with the integral VAR model.

REFERENCES

1. Bjorn-Helge Mevik and Ron Wehrens The pls Package: Principal Component and Partial Least Squares Regression in R, Journal of Statistical Software January 2007, Volume 18, Issue 2. http://www.jstatsoft.org/
2. Carreiro, Kapetorios & Marcellino (2010), Forecasting Government Bond Yields with large Bayesian VAR´s. Working papers No. 662 School of Economics and Finance.
3. Esposito Vinzi, V., W.W. Chin, J. Henseler and H. Wang (2007), Handbook of
    a. Partial Least Squares, Berlin: Springer.
4. Philip Hans Franses(2006) Forecasting 1 to h steps ahead using partial least squares
5. Econometric Institute, Erasmus University Rotterdam, Econometric Institute Report 2006-47
6. P.H. Garthwaite (1994). An Interpretation of Partial Least Squares. JASA Vol 89, No 425, pp122-127
    a. Hoskuldsson (1988). PLS Regression Methods. Journal of Chemometrics, Vol 2, pp 221-228.
7. Kim, J.H. (2001), Bootstrap after bootstrap prediction intervals for autoregressive models, Journal of Business & Economic Statistics, 19(1), 117-128.
8. Lutkepolh, H. (1991), Introduction to Multiple Time Series Analysis, 2nd ed., Springer-
    a. Verlag, Berlin.
9. Lutkepolh, H. (2006), Forecasting with VARMA models, in Elliot, G., C.W.J. Granger and
10. Timmerman (eds.), Handbook of Economic Forecasting, Vol. 1, 287-325.
11. McIntosh A.R., Bookstein F.L., Haxby J.V.and Grady C.L. (1996). Spatial Pattern Analysis of Functional Brain Images Using Partial Least Squares NeuroImage, Volume 3, Number 3, pp. 143-157(15). Academic Press
12. Marcellino, Stock & Watson (2004), A Comparison of Direct and Iterated Multistep AR Methods for Forecasting Macroeconomic Time Series.NBER
13. Paul Nomikos and John F. MacGregor (1995) Multivariate SPC Charts for Monitoring Batch Processes *Technometrics* Vol. 37, No. 1, pp. 41-59
14. Pascual, Ruíz y Fresoli (2011). *Bootstrap forecast of multivariate VAR models without using the backward representation*. Working Paper 11-34, Statistics and Econometrics Series,
15. Pascual, L., J. Romo, and E. Ruiz (2004a). Bootstrap predictive inference for ARIMA processes, Journal of Time Series Analysis, 25, 449-465
16. Pesaran, H. H., Pick, A., Timmermann, A. (*2010*). Variable Selection and Inference for Multi-period Forecasting Problems. Journal of Econometrics
17. Tiao, G.C., and D. Xu (1993), Robustness of maximum likelihood estimates for multi-step predictors: the exponential case, Biometrika, 80, 623-641

Acknowledgments: Partially supported by CONACYT CB No.105657 ; ECO2010-19357 Spain and FOMIX NL C42 -178237