

## Outline

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

## Motivation

- PLS es una técnica que ha mostrado su utilidad en muchas áreas de aplicación, tales como el control de procesos en la industria química en por ejemplo producción por batches, en imágenes médicas en donde se introducen modelos espacio temporales PLS, Path Modeling, Classification, Discrimination, Microarrays, just to mention some, with authors such as McGregor, Nomikos, MacIntosh, V. Esposito Vinz, Paul 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 standard assumptions are met
- Phillip Hans Franses (2006) propuso una metodología para realizar pronósticos  $h$  pasos adelante de manera óptima a través de una representación autorregresiva de orden  $p$ . “An Autorregresivo Parcial Least Square (PLS)” que denotamos como  $PLSAR(h, p)$

---

## Our case of interest

- Desarrollar un modelo para predecir la inflación en México
- El modelo considerará, cómo fuente principal de la dinámica inflacionaria mexicana, el crecimiento o variación de las condiciones monetarias del país
- Existe una gran discusión, incluso hoy en día, sobre si existe una relación de largo plazo entre el fenómeno monetario y el traspaso inflacionario que tiene éste
- No obstante, tiende a haber un consenso de que la inflación, en el largo plazo, es un fenómeno netamente monetario
- En este trabajo no abordaremos tal discusión, sino que se mostrarán las propiedades empíricas del modelo desarrollado en términos del error de predicción fuera de muestra

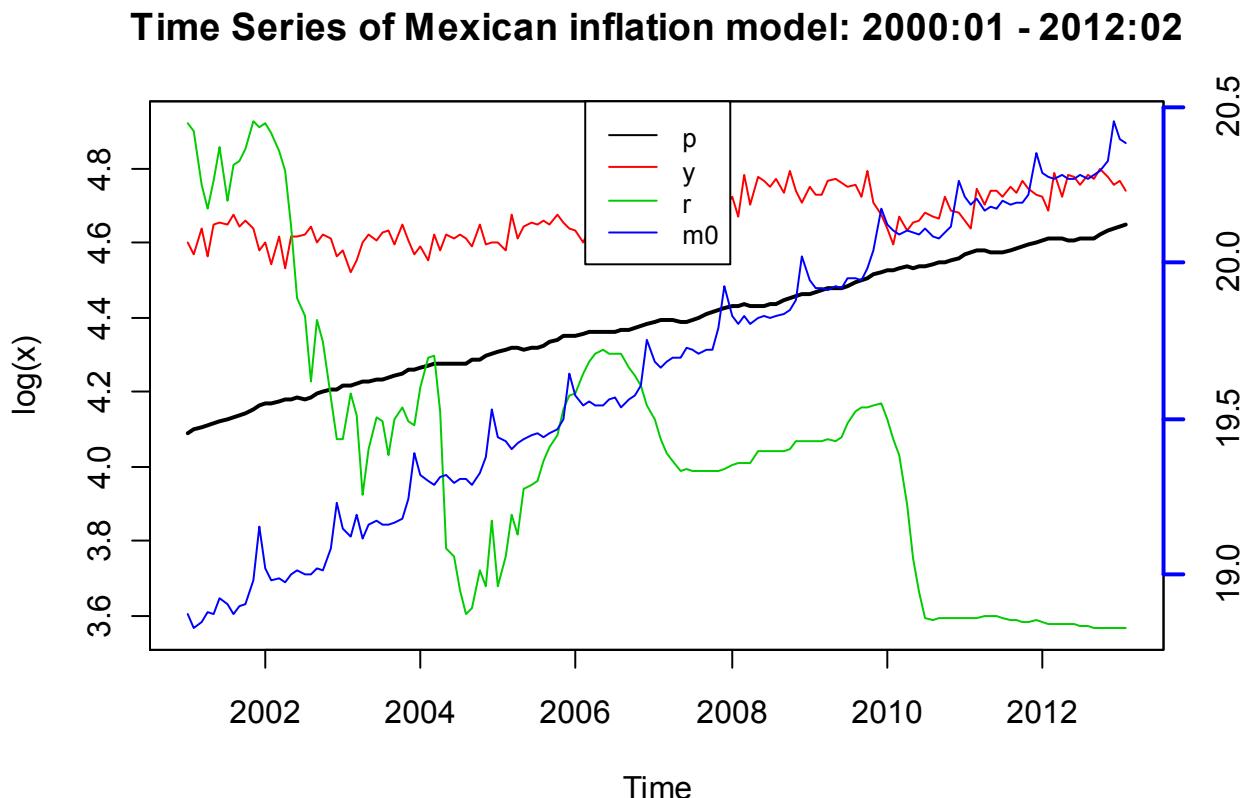
Se utilizan las siguientes variables de enero de 2000 a febrero de 2012:

- **p**: Consumer price index
- **m0**: Monetary 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.

- 
- Nosotros generalizamos la propuesta de Franses para lo que llamaremos  $VAR - PLS(h, p)$ , y se aplicará al modelo en cuestión, aunque solo mostraremos resultados comparativos
    1. Extensión multivariada dada la flexibilidad de los modelos Vectores Autorregresivos ( $VAR$ ), generando un modelo denominado  $VAR-PLS$
    2. Introducir a este concepto variables determinísticas (dummies, tendencias, etc.) y exógenas
    3. Construcción de intervalos de predicción vía Bootstrap

4. Construir un modelo VAR con capacidad predictiva y asimismo pronosticar con el VAR-PLS observando el grado de competitividad que tiene éste respecto a un modelo enfocado a predecir.
- 

$PLSAR(h,p)$

Frances 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)

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

- ❖ 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)

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

- ❖ El  $AR_h(p)$  es una alternativa al caso anterior, esto debido a que OLS minimiza la suma cuadrada de  $\hat{\varepsilon}_t$  pero no garantiza que sea mínima para  $h$  errores a futuro;
- ❖ Uno podría de contar con diferentes modelos para cada uno de los diferentes horizontes de pronóstico
- ❖ Para series de tiempo estacionarias, recordemos que el pronóstico de un  $AR(p)$  converge rápidamente a la media incondicional (obviamente la rapidez depende directamente de  $h \geq p$ )

Note: Para más detalles acerca de la literatura de este tipo de modelos, se puede consultar a Pesaran & Pick (2010), Marcellino, Stock & Watson (2004), Carreiro, Kapetorios & Marcellino (2010), Tiao & Xu (1993) entre muchos otros.

- 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

$$PLSAR(h,p): \hat{Y} = XB_{PLS}$$

- ❖ Es claro que existen correlaciones adyacentes entre las series de tiempo, la cual estos dos modelos iniciales no explotan. Dicho en otras palabras, se conoce que  $y_t$  y  $y_{t-j}$  están correlacionadas, más aún  $y_{T+h}$  y  $y_{T+h-j}$  también lo están, de esta manera, una alternativa viable es predecir conjuntamente  $(y_{t+h}, y_{t+h-1}, y_{t+h-2}, \dots, y_{t+1})$  con la ayuda de  $(y_t, y_{t-1}, y_{t-2}, \dots, y_{t-p})$ , y PLS una técnica atractiva para ello
- ❖ Frances propone reorganizar la información

- $(y_t, y_{t-1}, y_{t-2}, \dots, y_{t-p})$  en una matriz X (predictores)
- $(y_{t+h}, y_{t+h-1}, y_{t+h-2}, \dots, y_{t+1})$  en Y (predichos),
- y tomando la regresión a través de PLS, de tal manera que el proceso de construcción de variables latentes y cargas asociadas contengan la información relevante que tiene X en Y

En su trabajo muestra que un modelo PLS planteado como un fenómeno autorregresivo es competitivo respecto a modelos clásicos existentes en la literatura

## PLS

- El método de PLS puede ver desde diferentes ópticas, desde la perspectiva de interés en este trabajo, podemos relacionar la expresión del Vector Autorregresivo con la forma de un modelo lineal de la forma clásica:

$$Y = XB + U$$

donde  $Y$  es  $N \times k$ ,  $X$  es una matriz  $n \times N$ ,  $B$  es una matriz de  $(N + 1) \times k$  y  $U$  es una matriz de  $N \times k$ .

- Bajo la concepción de Franses, de explotar las correlaciones existentes entre  $(y_t, y_{t-1}, y_{t-2}, \dots, y_{t-p})$  y  $(y_{t+h}, y_{t+h-1}, y_{t+h-2}, \dots, y_{t+1})$ , la diferencia escencial entre el VAR y PLS radicará únicamente que mientras en el primero usa directamente OLS, el segundo usa proyecciones sobre variables latentes, esto es, una descomposición que maximiza la covarianza existente una matriz  $X = Y'_{t-1}$  con  $Y = Y'_t$

The basic procedure maximize the  $\text{cov}(X\alpha, Y\beta)^2$  under

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

where

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

- $(X\alpha, Y\beta)$  representa una combinación lineal de las variables que maximizan la covarianza o covarianza al cuadrado (no interesa el signo sino la maximización)
- $S_{xx}$  y  $S_{yy}$  the varince covariance matrices;  $\beta'\beta = 1$  &  $\alpha'S_{xx}\alpha = 1$

maximizando nuestra función objetivo, obtenemos:

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

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

normalizando los scores  $t = t/\sqrt{t't}$  y continuando con más simplificaciones se obtienen los loadings para  $X$  y  $Y$ :  $p = E't$  y  $q = F't$ . Agrupamos cada  $w, t, p$  y  $q$  en matrices  $R = W(P'W)^{-1}$  finalmente se reconstruye

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

Note: For a nice introduction see : P.H. Garthwaite (1994). An Interpretation of Partial Least Squares. JASA Vol 89, No 425, pp 122-127.

Agnar Hoskuldsson (1988). PLS Regression Methods. Journal of Chemometrics, Vol 2, pp 221-228.

## Vectores Autorregresivos y Mínimos Cuadrados Parciales

Un proceso  $VAR(p)$  se define como:

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

$A_i$ : Matrices de coeficientes para  $i = 1, 2, \dots, p$

$\varepsilon_t$ : Proceso ruido blanco con matriz de covarianza positiva definida  $E[u_t u_t'] = \Sigma_u$

$C$ : Matriz de potenciales regresores determinísticos

$D_t$ : Vector columna de regresores determinísticos apropiados

Notemos que el  $VAR(p)$  puede ser definido como un  $VAR(1)$  de la forma

$$Y_t = AY_{t-1} + V_t \quad (5)$$

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

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

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

- Si los valores propios de  $A$  son menores que uno, entonces el proceso  $VAR(p)$  es estable.

Una parte importante de la especificación del modelo está dada por encontrar el orden  $p$  del  $VAR$  a estimar. El procedimiento general consiste en ordenar los  $p = 0, \dots, p_{max}$  y elegir el valor de  $p$  que minimiza algún criterio de selección. La selección del criterio para el  $VAR(p)$  tienen la forma

$$IC(p) = \ln|\hat{\Sigma}(p)| + C_T \cdot \varphi(K, p) \quad (6)$$

Donde  $\hat{\Sigma}(p) = T^{-1} \sum_{t=1}^T \hat{u}_t' \hat{u}_t$ ,  $C_T$  es una secuencia indexada por el número de realizaciones de  $T$  y  $\varphi(K, p)$  es la función de penalización por la longitud del modelo  $VAR(p)$ . Los cuatro criterios de información más utilizados son los criterios de Akaike (AIC), Schwarz-Bayesiano (BIC), Hannan-Quinn (HQ) y el Error Final de Predicción (FPE):

$$AIC(p) = |\hat{\Sigma}(p)| + \frac{2}{t} p K^2 \quad (7a)$$

$$BIC(p) = |\hat{\Sigma}(p)| + \frac{\ln T}{t} p K^2 \quad (7b)$$

$$HQ(p) = |\hat{\Sigma}(p)| + \frac{2\ln \ln T}{t} p K^2 \quad (7c)$$

$$FPE(p) = \left( \frac{T+p^*}{T-p^*} \right)^K \det(\hat{\Sigma}(p)) \quad (7d)$$

- El criterio AIC sobreestima asintóticamente el orden con probabilidad positiva, mientras que el BIC y HQ estima consistentemente el orden bajo ciertas condiciones tales como, si el verdadero valor de  $p$  es menor o igual que  $p_{max}$

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

$$y_{T+h|T} = A_1 y_{T+h-1} + \dots + A_p y_{T+h-p} + CD_{T+h} \quad (8)$$

- La estimación de  $A_i$  es generalmente llevada a cabo por OLS

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

- Bajo ciertas condiciones de generalidad de comportamiento estacionario y ergodicidad en los modelos VAR (Hamilton (1994), Lutkepohl (1991) entre otros),  $\text{vec}(\hat{A})$  es consistente y distribuido asintóticamente con matriz de covarianza:

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

Donde

- $\hat{\Sigma} = \frac{\sum_{t=1}^T \hat{\varepsilon}_t' \hat{\varepsilon}_t}{T-K}$

- $\hat{\varepsilon}_t = Y_t - \hat{A}'Z_t = Y_t - \hat{A}'Y_{t-1}$  es el residual de mínimos cuadrados en el tiempo  $t$
- El  $i$ -ésimo elemento de  $\text{vec}(\hat{A})$  es asintóticamente normal (para un VAR estable) con errores estándar dados por la raíz cuadrada de los elementos de la diagonal de  $\hat{\Sigma} \otimes (Z'Z)^{-1}$ .
- Las pruebas  $t$  son válidas asintóticamente para los coeficientes estimados.

Una situación de interés radica en la presencia de una o más raíces unitarias en las  $y_j$ . Esto ha dado pie a toda una teoría económica, que se basa en modelar el comportamiento de largo plazo y analizar la dinámica temporal de un conjunto de series.

- **Cointegración:** Las componentes del vector  $y_t$  se dice que son cointegradas de orden  $d, b$ , denotado por  $y_t \sim CI(d, b)$  si
  - todos los componentes de  $y_t$  son  $I(d)$ ; y
  - existe un vector  $\beta (\neq 0)$  tal que  $z_t = \beta'y_t \sim I(d-b), b > 0$ . El vector  $\beta$  es llamado vector de cointegración
- **Corrección de error:** El vector bivariado  $y_t = (y_{1t}, y_{2t})'$  con vector de cointegración  $\beta = (1, -\beta_2)'$ , entonces  $\beta'y_t = y_{1t} - \beta_2 y_{2t} \sim I(0)$  existe una representación de corrección de error de
  - $\Delta y_{1t} = \alpha_1 + \gamma_1(y_{1t-1} - \beta_2 y_{2t-1}) + \sum_{i=1}^K \psi_{1,i} \Delta y_{1t-i} + \sum_{i=1}^K \psi_{2,i} \Delta y_{2t-i} + u_{1t}$
  - $\Delta y_{2t} = \alpha_2 + \gamma_2(y_{1t-1} - \beta_2 y_{2t-1}) + \sum_{i=1}^L \xi_{1,i} \Delta y_{1t-i} + \sum_{i=1}^L \xi_{2,i} \Delta y_{2t-i} + u_{2t}$

Notemos que el modelo  $VAR(p)$  puede expresarse como:

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

$$\begin{aligned}\Gamma_i &= -(A_{i+1} + \dots + A_p), \quad i = 1, \dots, p-1 \\ \Pi &= -(I - A_1 - A_2 - \dots - A_p)\end{aligned}$$

Modelo de Vector de Corrección de Error (VECM) transitorio

O bien :

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

$$\begin{aligned}\Gamma_i &= -(I - A_1 - A_2 - \dots - A_i), \quad i = 1, \dots, p-1 \\ \Pi &= -(I - A_1 - A_2 - \dots - A_p)\end{aligned}$$

VECM de largo plazo

De esta manera, la matriz  $\Pi$  tiene los siguientes puntos de interés:

1.  $\text{rk}(\Pi) = n$ , todas las  $n$  combinaciones deben ser estacionarias; el VECM es un modelo VAR en niveles
2.  $\text{rk}(\Pi) = 0$ , no existe una combinación lineal estacionaria tal que  $\Pi y_{t-1}$  sea estacionaria, excepto la solución trivial, i.e. es un modelo  $VAR(p-1)$  en primeras diferencias

3.  $0 < rk(\Pi) < n$ , en este caso  $\Pi = \alpha\beta'$  con dimensión  $(nxr)$  y  $\beta'y_{t-1}$  es estacionaria. Cada columna de  $\beta$  representa una relación de largo plazo
- Si el objetivo es pronosticar, aún en el caso de variables integradas y cointegradas, hacerlo mediante la representación VAR es muy apropiado (ver Lutkepohl 2006)
- 

From the application:

- Se especificó el orden  $p$  del VAR teórico a través del criterio de Error Final de Predicción el cuál fue de 12, asimismo, se realizó la prueba de Johansen para denotar la presencia o no de relaciones de largo plazo.
- Los resultados obtenidos mostraron que al 1%, 5% y 10% existe una ecuación de cointegración, la cual está dada por la siguiente expresión:

$$p + 9.68 - 0.43m - 0.89y + 0.10r = 0 \quad \boxed{=}$$

- Lo anterior es congruente con la realidad ya que establece el traspaso inflacionario está impulsado por el crecimiento monetario, exceso de demanda y la reducción del costo del dinero.

## VAR-PLS

1. Se propone utilizar la representación de un modelo VAR, de tal manera que sea el proceso generador de datos
2. El modelo VAR, será quien dictamine el orden del proceso autorregresivo que será utilizado en la regresión PLS
3. Trabajaremos con matrices en estructura de rezagos según cada una de las variables

- Agrupamos en la matriz  $X$  la información observada en el pasado como en

$$X = Y_{t-1} = \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p} \end{bmatrix} \quad y$$

- En  $Y$  lo observado en tiempo  $t$   $Y = Y_t = \begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{bmatrix}$  estimando el modelo a través del proceso de construcción de variables latentes.

- It is simple to introduce variables exógenas o determinísticas through a matrix  $C$ ,

modificando la composición de  $X = Y_{t-1}^* = \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p} \\ D_t \end{bmatrix}$  y por ende de la matriz de

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

- De esta forma obtenemos un modelo VAR-PLSX con la única finalidad de forecast  $h$  steps ahead

Como el objetivo es construir un modelo de pronóstico robusto que a su vez sea comparado con la técnica de VAR-PLS se realiza el siguiente procedimiento

Es importante señalar que, para el VAR-PLS podemos obtener hasta  $pK$  componentes. Para este ejercicio se estiman cada una de ellas sólo con fines comparativos, y para observar la capacidad predictiva respecto a métodos recursivos de pronóstico construidos a partir de un modelo VAR

- Se **guardaron** las 24 observaciones finales con el fin de obtener una ventana de tiempo de horizonte largo que sirva para comparar cada uno de los modelos
- Para determinar el orden del VAR-PLS, primero se determina la selección del rezago óptimo a través de algún criterio de información.
- Para verificar que tal proceso generador de datos sea consistente con la teoría económica, se estima el  $VAR(p)$  denotando su característica estocástica, es decir, si es cointegrado o no, obteniendo en su caso, los coeficientes de largo plazo.

- Para ese  $p$  óptimo con una especificación constante y sin variables determinísticas, se procede a estimar el VAR-PLS con  $h = 24$ , obteniendo el error fuera de muestra para cada una de las componentes existentes  $pK = 48$
- Para el modelo  $VAR(p)$ , se tienen 4 variables, 4 posibles especificaciones del VAR (ninguna, constante, tendencia, constante y tendencia), 11 posibles variables dummies estacionales (una para cada mes). Combinado todas de las variables, con las especificaciones y variables determinísticas posibles nos da un total de  $VAR_j(p)$   $j = 1, \dots, 484$  posibles a estimar.
- Se estiman cada uno de los modelos y se seleccionan los que minimizaron 7 criterios de error fuera de muestra, es decir, para esto se eliminaron otras  $h = 24$  observaciones, con el fin computar los estadísticos de error (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
- Adicionalmente se trabajó con un modelo benchmark (Autorregresivo de orden 1) para el cual se calcularon los  $i = 1, \dots, h$  ( $h = 24$ ) pronósticos generando el siguiente estadístico:

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

Posteriormente los siguientes 3 estadísticos:

- MRAE: Mean Relative Absolute Error
- MdRAE: Median Relative Absolute Error
- GMRAE: Geometric Mean Absolute Error
- De esta manera se obtienen 7 modelos, los cuales se integran en un solo número obtenido del cuantil 50% para cada uno de los horizontes de pronóstico.
- Se estima el  $VAR(12) - PLS(24, j)$  y se realizan los  $pK = 48$  modelos, obteniendo para cada serie el *MAPE* para fines comparativos para un horizonte de pronóstico de 24 datos.

### Intervalo de predicción: VAR-PLS

- Uno de los objetivos de este trabajo radica en la construcción del intervalo de predicción para el modelo VAR-PLS

El procedimiento para su realización es similar a lo propuesto por Pascual, Ruíz y Fresoli (2011). *Bootstrap forecast of multivariate VAR models without using the backward representation*. Working Paper 11-34, Statistics and Econometrics Series, que a su vez utilizan las ideas seminales de Kim (2001) y de Pascual, L., J. Romo, and E. Ruiz (2004a). *Bootstrap predictive inference for ARIMA processes*, Journal of Time Series Analysis, 25, 449-465

- La idea cuando se realiza para el VAR es :
  - 1) Introducir la incertidumbre debida a la estimación de parámetros y corregir así las regiones de confianza asociadas (válidas bajo asintóticamente bajo la distribución Gaussiana)
  - 2) Simplificar computacionalmente el cálculo con VAR que tienen muchos rezagos y regularmente implican que la representación backwards resulte muy compleja
- El procedimiento debe de adecuarse según la construcción de las matrices de coeficientes de regresión a través de PLS, en este caso su representación VAR

El procedimiento se resume cómo:

1. Estimamos  $Y_t = X_t \hat{B}_{PLS} + \hat{U}_t$
2. Obtenemos  $\hat{U}_t^* = \hat{U}_t - \bar{\hat{U}}_t$  para obtener una muestra ordenada de tales residuos centrados y reescalados.
3. Con los  $p$  valores iniciales  $Y_0 = \{Y_p, \dots, Y_1\}$  y con lo obtenido en 1) y 2) generamos  $Y_t^*$  a través de:

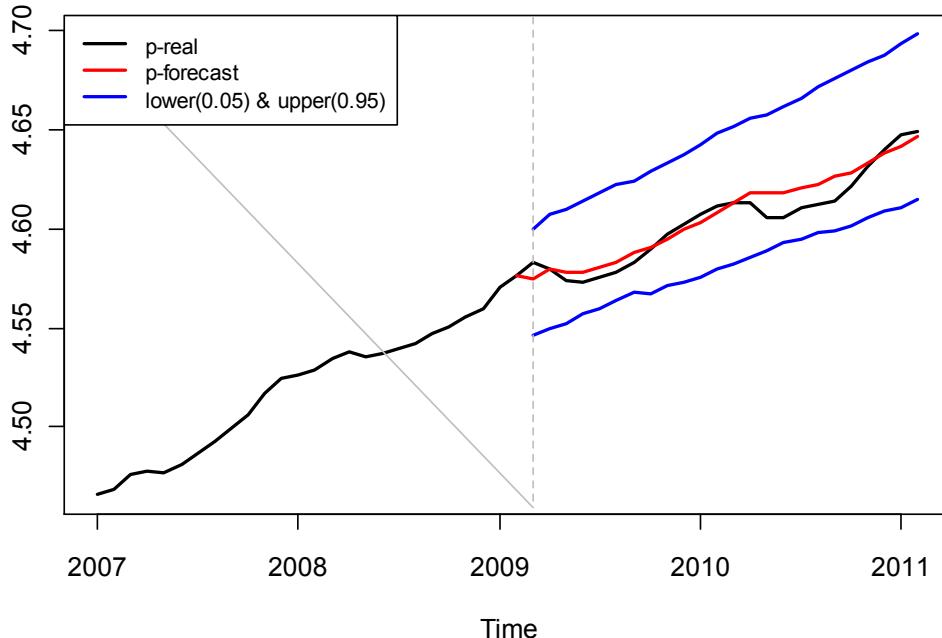
$$Y_t^* = X_t \hat{B}_{PLS} + \hat{U}_t^* \quad t = 1, \dots, n - p$$

4. Obtenemos  $\hat{Y}_{T+h}^*$  con repitiendo el paso 2 a 4 para  $n = 1, \dots, N$  ocasiones
5. Finalmente, para cada una de las  $n^{th}$  variables y los  $N$  conjunto de pronósticos obtenemos:

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

Donde  $q_B^*(\tau)$  es el  $\tau^{th}$  percentil de  $G_{n,B}^*(x) = \#(y_{n,T+k}^{*(b)} \leq x) / N$

### Forecast series p: VAR(12)-PLS(h=24,k=15)



- Podemos observar que prácticamente, el valor real es el mismo que el pronóstico obtenido con un error absoluto promedio de 0.11%. Se observan los intervalos de predicción vía Bootstrap que “atrapan” en cada uno de los casos el valor real de la serie.
- Para términos netamente económicos, ésta aproximación puede resultar competitiva

NOTA: Al ser un modelo multivariado se obtienen el resto de los pronósticos de las series, sin embargo dado que el objetivo es predecir la serie de precios, no se considera relevante para este trabajo observar a detalle el comportamiento, no obstante es válido comentar que pronosticó (con esta misma componente) **con un error de 0.23% para m0, para y 0.83% y finalmente con un 5.84% la tasa de interés.**

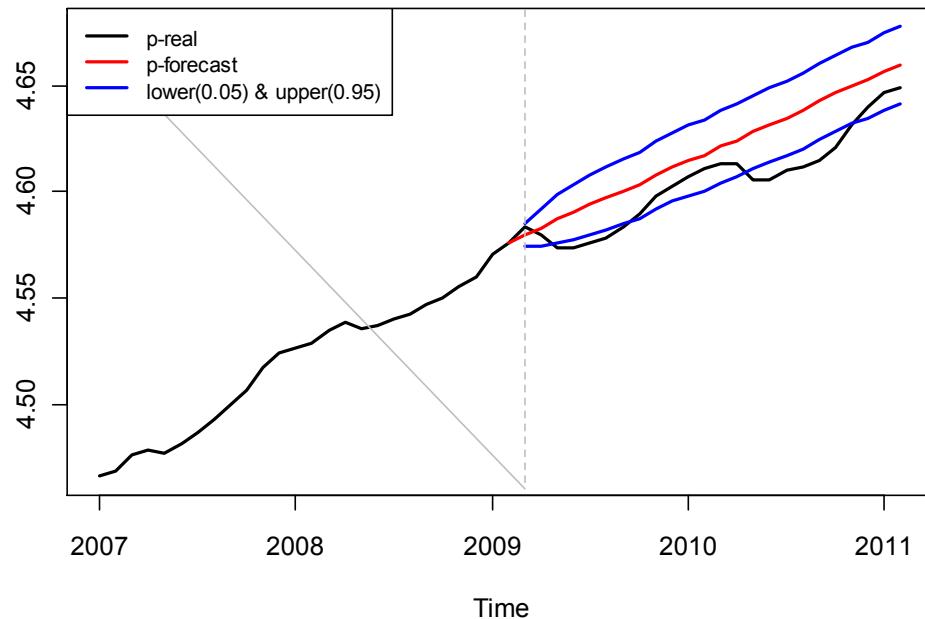
Para el Integral-VAR, los  $VAR_j(p)$  óptimos fueron los siguientes:

**[Cuadro 1:] Modelos VAR óptimos**

Criterio	MAPE	MdAPE	RMSPE	RMdSPE	MRAE	MdRAE	GMRAE
<b>Estadístico</b>	<b>0.29</b>	<b>0.30</b>	<b>0.33</b>	<b>0.30</b>	<b>0.24</b>	<b>0.22</b>	<b>0.18</b>
<b>Rezagos</b>	2	2	2	2	2	2	3
<b>Estacionalidad</b>	10	10	10	10	4	11	7
<b>Especificación</b>	Constante y tendencia						

Integrando los pronósticos y el intervalo de predicción con su respectivo cuantil del 50% para cada modelo y horizonte de pronóstico obtenemos lo siguiente:

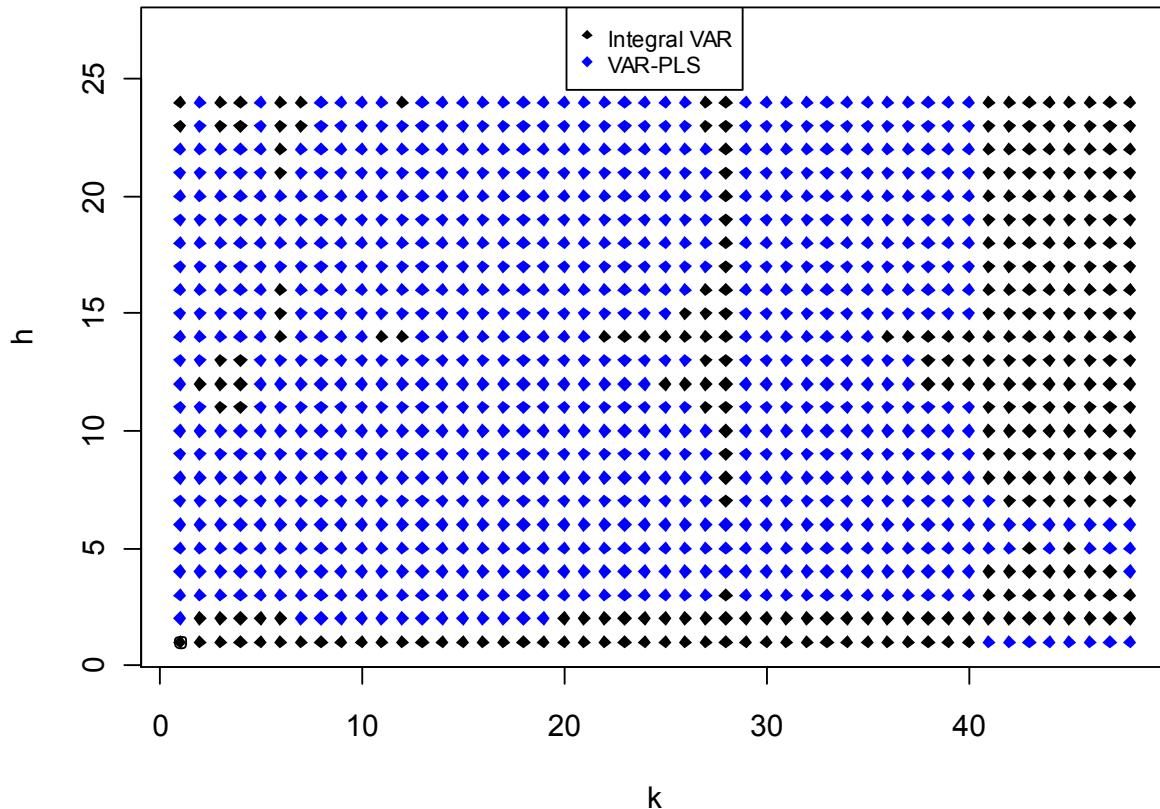
### Forecast series p: Integral-VAR



Aunque en términos numéricos el error fue de 0.32%, a detalle se observan grandes diferencias respecto al real observado, sobreestimando la inflación en cada uno de los casos. No obstante, se decidió realizarlo de esta manera para hacer comparativo las variables que intervienen en los pronósticos para ambas especificaciones.

El siguiente gráfico muestra cómo se comportó **cada componente respecto al Integral VAR**:

## Integral VAR vs VAR-PLS (Comparasion with MAPE)



Resulta muy interesante que en promedio, **un 73.52% de las veces, fueron superior las componentes** y hace sentido que las componentes más lejanas dejan de tener efecto, dado a que son éstas las que el tienen menor explicación de la variabilidad observada.

En otras palabras, la metodología VAR-PLS parece ser atractiva respecto su competidor inmediato, **que es un VAR enfocado a predecir**.

### Conclusiones

- En este trabajo se presentó una metodología alterna para realizar pronósticos multivariados mediante la perspectiva de explotar un conjunto de series de tiempo  $y_{t+j}$  con  $y_t$ , aprovechando la naturaleza que tiene por construcción la técnica de PLS, específicamente al momento de plantear un modelo lineal, situación que se presenta para el caso de un VAR.
- Se planteó un modelo que considera como principales fuentes inflacionarias un agregado monetario (billetes y monedas), tasas de interés y una variable que ingreso y se realizaron dos metodologías de pronóstico;

- la primera de ellas un VAR-PLS cuyo orden  $p$  está dado por el criterio de FPE de un VAR, estimando en este caso, todas y cada una de las  $pK$  componentes posibles.
- Finalmente, los intervalos de predicción se hicieron vía Bootstrap

Lo anterior representa un punto de partida que indica una posibilidad latente de que el VAR-PLS es una técnica atractiva de pronóstico multivariado

#### Líneas futuras a trabajar

- Cómo el incluir en el ejercicio de cointegración, la técnica de PLS, denotando las regularidades teóricas que esto implica,
- Construir modelos PLS-VAR que integren cada una de las componentes
- Poner en competencia la metodología VAR-PLS con VAR que [mezclen](#) las variables que intervienen dentro de un pronóstico.