Descomposición Factorial de la Inflación en Perú

Alberto Humala (BCRP)   Gabriel Rodríguez (BCRP)

XXVI Encuentro de Economistas
Banco Central de Reserva del Perú

26-28/11/2008

Motivación

- Análisis factorial permite reducir la dimensión y es una herramienta estadística útil
- Varios ejemplos económicos pueden ser naturalmente analizados bajo este enfoque:
  - La teoria de arbitraje de precios . La APT de Ross (1976) señala que un número reducido de factores puede ser usado para explicar un gran número de retornos de activos;
  - Predicción con indices de difusión. Stock y Watson (1998, 1999) predicen la inflación usando indices de difusión (*"factores"*) que son construidos a partir de un vasto número de series macroeconómicas. La hipótesis subyacente es que esas series son explicadas por un pequeño número de factores no observables

Objetivos

- Separación de la inflación de bienes de consumo en tres componentes: cambios en precios relativos idiosincráticos, cambios en precios relativos agregados y cambios en precios absolutos.
- Identificación de una medida de *"inflación pura"* definida como el componente común en la inflación de bienes que tiene un efecto equiproporcional sobre todos los precios y que no está correlacionada con cambios en precios relativos

Modelos Factoriales

- Considere Y_{it} los datos observados para la unidad i^{th} al periodo t, para $i = 1, 2, 3, ..., N$ (rubros de la canadasta) y $t = 1, 2, 3, ..., T$ (observaciones)
- Sea el modelo siguiente:

$$Y_{it} = \lambda_i' F_t + u_{it}, \tag{1}$$

donde:

  - F_t es el vector de factores comunes;
  - λ_i es el vector de coeficientes asociados a F_t;
  - u_{it} es el componente idiosincrático de Y_{it}.

El Modelo (1)

- Denote la tasa de cambio para el rubro i^{th} entre t y $t-1$ como π_{it}
- Sea π_t un vector $N \times 1$ que contiene los precios de los N rubros considerados
- El co-movimiento entre los precios son modelados usando el modelo lineal factorial siguiente:

$$\pi_t = \Lambda F_t + u_t, \tag{2}$$

donde:

  - F_t es un vector $(k \times 1)$ conteniendo los k factores que explican las fuentes comunes de variación en precios;
  - Λ es una matrix $(N \times k)$ que contiene los coeficientes que indican cómo el precio del rubro i^{th} responde a los choques;
  - u_t es un vector $(N \times 1)$ que captura la variabilidad de precios relativa asociada con eventos sectoriales idiosincráticos o errores de medida.

El Modelo (2)

- Factores en F_t explican una porción importante de la variación en los datos
- Cambios en los precios de los rubros proveen información sobre los choques agregados
- Este componente agregado se separa en un componente de cambio en precios absolutos (escalar denotado por a_t) y potencialmente varios componentes de precios relativos (vector de tamaño $k-1$ denotado por R_t). Esta descomposición puede ser escrita

$$\Lambda F_t = \beta a_t + \Gamma R_t \tag{3}$$

  donde
  - β es un vector $N \times 1$ de unos (cambios en precios absolutos cambian todos los precios equiproporcionalmente)
  - Γ es una matriz $N \times (k-1)$ (cambios en precios relativos afectan precios en diferentes proporciones)
- Pregunta: Las fuerzas comunes de variación (ΛF_t) puede ser descompuesta?

El Modelo (3)

- Dos inconvenientes:
  - El vector β puede estar ausente del espacio de Λ (no hay cambios en precios absolutos en los datos);
  - La descomposición en (3) no es única. Es decir: a_t y R_t no pueden ser identificados separadamente. Intuición: los cambios absolutos en precios no pueden ser distinguidos de un cambio en "precios relativos promedio". Hay muchas maneras de definir este promedio.
- Solución: definir la noción de *"inflación pura"*

$$v_t = a_t - E[a_t|\{R_\tau\}_{\tau=1}^{T}] \tag{4}$$

- Interpretación: v_t es el componente común en cambios en los precios que tiene un efecto equiproporcional sobre todos los precios y que no está correlacionado con los cambios en precios relativos

Estimación (1)

- Tres etapas:
  - Estimar el número de factores (k);
  - Estimación de los factores a_t y R_t. Asimismo estimación de la matriz Γ y examinar la restricción unitaria en β;
  - Estimar v_t
- Análisis de robustez
- Para la elección del tamaño k del vector F_t seguimos Bai and Ng (2002).
  - El método permite estimar k de manera consistente cuando $min(N, T) \Rightarrow \infty$;
  - Método basado en el número de valores propios dominantes de la matriz de covarianza (o correlación) de los datos;
  - Complemento: análisis de aporte marginal de factores adicionales

Estimación (2)

- Estimación de (1)-(2) por componentes principales restrictos. Esto significa resolver el problema de mínimos cuadrados restringidos siguiente:

$$\min_{\beta,\Gamma,a,R} \sum_{i=1}^{N} \sum_{t=1}^{T} (\pi_{it} - \beta_i' a_t - \gamma_i' R_t)^2 \qquad (5)$$

- Estimación paramétrica con algunos supuestos sobre el comportamiento de las variables latentes a_t, R_t y u_{it}. Estimación por Máxima Verosimilitud:
  - Se asume que (a_t, R_t) sigue un proceso VAR
  - El proceso u_{it} sigue un proceso AR independiente
  - Todos los errores son asumidos Normales.

Estimación (3)

- El modelo de componentes no observados es

$$\pi_{it} = \beta_i' a_t + \gamma_i' R_t + u_{it}, \tag{6}$$

$$\Phi(L) \begin{pmatrix} a_t \\ R_t \end{pmatrix} = \epsilon_t, \tag{7}$$

$$\rho_i(L) u_{it} = \alpha_i + e_{it}, \tag{8}$$

donde

- $\{e_{it}\} \sim N(0, \sigma_i^2)$, $\{e_{jt}\} \sim N(0, \sigma_j^2)$, $\{\epsilon_t\} \sim N(0, Q)$ mutuamente y serialmente no correlacionados.
- Columnas de Γ son normalizadas para identificar los factores. Normalización: las columnas de Γ son mutuamente ortogonales.

Estimación (4)

- Tamaño del modelo es grande. Maximización numérica de la función de verosimilitud es computacionalmente compleja;
- Ejemplo. En nuestro caso $N = 45$. Asumiendo un VAR(4) para (7) y un AR(1) para u_{it} y tres factores ($k = 3$) tenemos un total de 261 parámetros a ser estimados;
- Estructura lineal de las variables latentes permite usar el algoritmo: EM (filtro de Kalman (E) y regresión lineal (M);
- Los estimados de $\Phi(L)$ permiten calcular la proyección implicada en (4) para obtener v_t.

Datos

- Información mensual de precios a nivel de rubros: 1995:01-2008:07
- Inflación calculada como $1200 \times [\ln(P_{it}\text{-}\ln P_{it-1}]$
- Tratamiento de observaciones atípicas
- N final: 45 rubros.

Resultados (1)

- Presencia de dos factores: a_t y R_t
- La hipótesis nula de$\beta = 1$ no puede ser rechazada
- Resultados robustos al cambio del orden del VAR(p). Se probó con $p = 4, 6, 8$
- Resultados robustos a la especificación del comportamiento estocástico de a_t y R_t. Se usaron combinaciones en las cuales $a_t \sim I(1)$ y $R_t \sim I(0)$ o viceversa o ambos casos
- El estimado de v_t (*inflación pura*) sigue un comportamiento similar al de la inflación subyacente
- Estimados de v_t son más volátiles usando el enfoque de componentes principales comparado con el de Máxima Vesosimilitud

Resultados (2)

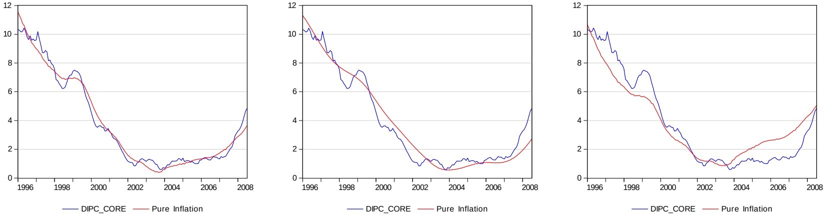

Figura 1a. $a_t \sim I(1)$, $Rt \sim I(0)$; Figure1b: $a_t \sim I(1)$, $R_t \sim (1)$; Figure 1c: $a_t \sim I(0)$, $R_t \sim I(0)$

Conclusiones

- Información desagregada a nivel de precios o inflación puede ser reducida a unos cuantos factores con interpretación económica;
- La medida de inflación obtenida es comparable a la inflación subyacente;
- Los resultados muestran robustez a diferentes especificaciones respecto de los rezagos usados en las estimaciones asi como respecto al comportamiento estocástico de los factores estimados