

PROYECCIONES DE INFLACIÓN CON MODELOS SPARSE VAR ROBUSTOS

Carlos Barrera Chaupis, DME*

Resumen

La idea general de robustez de un estimador se refiere a la menor sensibilidad del mismo ante la presencia de un porcentaje relativamente pequeño de observaciones fuera de patrón previamente establecido, llamadas también *outliers*. El objetivo principal del presente trabajo es utilizar estimadores robustos en modelos SparseVAR que a su vez generen proyecciones robustas de la inflación del Índice de Precios al Consumidor (IPC), es decir, menos sensibles a la presencia de *outliers* a la vez que se mantiene o incluso mejora su precisión.

Existe una gama extensa de estimadores robustos para modelos lineales estáticos uni-ecuacionales, una gama más limitada para modelos lineales estáticos multi-ecuacionales y, en general, muy pocos procedimientos robustos para modelos lineales dinámicos. Un primer objetivo práctico del presente estudio es reseñar las estrategias de estimación robusta (procedimientos robustos) más importantes, poniendo énfasis en aquellos procedimientos que serán aplicables en modelos SparseVAR.

El procedimiento propuesto parte de la obtención de restricciones de exclusión utilizando criterios de información a partir de estimadores robustos uni-ecuacionales que puedan ser utilizados con los limitados recursos computacionales disponibles. Una vez obtenidas estas restricciones, inherentes a modelos VAR parsimoniosos, el conjunto del sistema es estimado mediante estimadores robustos multi-ecuacionales, para luego generar sus proyecciones.

El segundo objetivo práctico del estudio es comparar la precisión *ex post* de estas proyecciones de inflación elaboradas con estimadores menos sensibles a *outliers* y la de las proyecciones basadas en estimadores que se conoce son sensibles a ellas, manteniendo siempre el principio de parsimonia inherente a los modelos Sparse VAR.

0. INTRODUCCIÓN

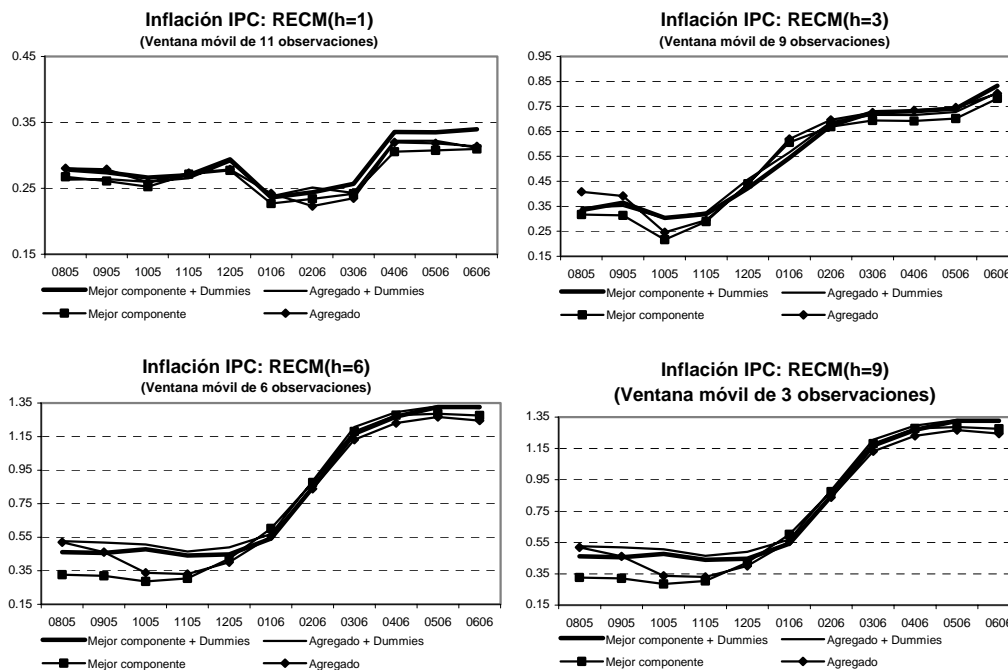
Los modelos Sparse VAR han tenido un relativo éxito en términos de una tendencia creciente en la precisión de sus proyecciones de inflación, como se ha reseñado previamente (Barrera (2005, 2006)). Aunque los niveles de precisión habían mejorado sobremanera hasta alrededor de setiembre del 2005, las proyecciones de estos modelos son sensibles ante la eventual presencia de observaciones fuera del patrón previamente establecido¹, lo que implica una eventual reducción en su precisión. Por ejemplo, la secuencia de *outliers* en el IPC desde el último trimestre del 2005 hasta alrededor de mayo del 2006 redujo abruptamente la precisión de las proyecciones, para luego elevarse nuevamente, aunque más lentamente, a medida que se ingresan nuevas observaciones dentro del patrón. El Gráfico 1 muestra la precisión de cuatro

* El autor agradece los comentarios realizados en el Seminario de Investigación Económica organizado por la Subgerencia de Investigación Económica, en particular los de Paul Castillo, Alberto Humala, Zenón Quispe, Vicente Tuesta y Marco Vega. Todo error remanente es responsabilidad del autor.

¹ Nótese además que si existen *outliers* al interior de la muestra, debe ser evidente que el "patrón" no puede estar fuertemente establecido.

modelos Sparse VAR utilizados para proyectar la inflación IPC, específicamente para los horizontes $h \in \{1, 3, 6, 9\}$, cuyo cálculo considera el mismo número de errores de predicción ejecutados (observaciones) para permitir la comparación a medida que aparece la secuencia de *outliers* mencionada. Las unidades son siempre variaciones porcentuales 12 meses en el IPC mensual².

Gráfico 1
Inflación IPC : Headline



Existen dos diagnósticos paralelos para este abrupto aumento en las raíces del error cuadrático medio (RECM). El primer diagnóstico es honesto al aceptar que se mantienen problemas de especificación importantes en el modelo, los cuales dificultan capturar correctamente el patrón que las "observaciones fuera de patrón" pueden estar señalando. La propia literatura sobre modelos VAR suele aceptar problemas como la omisión de variables importantes (conjuntos de información todavía limitados), factores idiosincráticos o temporales no observables usualmente presentes en bases de datos desagregadas, así como en componentes no lineales importantes (evidentemente omitidos en modelos lineales). Asimismo, en el caso de modelos dinámicos, las variables presentes pueden estar sujetas a choques temporales potencialmente asociados a estos problemas y que pueden presentarse en diferentes tramos de la muestra histórica. Bajo este diagnóstico, surgen errores o choques fuera de cierto intervalo por lo que la evaluación de la precisión *ex post* de las proyecciones de cualquier modelo se debe realizar con los datos históricos "duros", que no requieren de corrección.

² Otro ejemplo reciente de deterioro eventual en la precisión por la ocurrencia de *outliers* en la inflación IPC es la secuencia de *outliers* ocurrida durante los primeros 9 meses de 1998 y asociada al fenómeno de "El Niño". Aunque no se ha documentado esta reducción en la precisión de las proyecciones, resultó una fuente de motivación para el presente trabajo el hecho de que la inflación IPC del año 1998 fuera predicha correctamente con una combinación de modelos VAR bivariados estimados con datos a diciembre 1997.

También es posible que los propios datos estadísticos sean "medidos directamente" mediante encuestas que adolecen de una cobertura limitada por falta de recursos o de un diseño estadístico que induce una sensibilidad inherente o, alternativamente, que los datos se obtengan como una combinación de datos directamente medidos con datos medidos indirectamente, como suele ser el caso del P.B.I., donde las medidas indirectas adolezcan de coeficientes con elevada incertidumbre o incluso *ad hoc*, etc. Bajo este segundo diagnóstico, los eventos fuera del patrón, y por ende, los errores o choques fuera de cierto intervalo, reflejan estos "problemas de construcción" de los propios datos.

En ambos casos, las observaciones fuera del patrón previamente establecido mantienen una presencia muestral importante reflejándose en un número de errores fuera de cierto intervalo que resulta mayor al que se obtendría de multiplicar la probabilidad que una distribución Normal asigna a observaciones o errores extremos (distribución con "colas" delgadas) por el tamaño de la muestra. En ambos casos, los coeficientes obtenidos con el estimador MCO-MV-Normalidad serán extremadamente sensibles a estos eventos fuera del patrón, es decir, el estimador por MCO-MV-Normalidad será "poco robusto".

Como la presencia de *outliers* se da en toda la muestra, la solución es utilizar estimadores robustos basados en distribuciones con "colas" anchas o de la detección y eliminación de aquellas observaciones fuera de patrón presentes en la muestra disponible. En este contexto parecería natural afirmar que una "evaluación de la precisión *ex post* de las proyecciones" requiere de una muestra de datos "corregida", por lo que los correspondientes RECMs serían menores a los calculados sin esta corrección de los datos históricos iniciales. Sin embargo, esta forma de "evaluación" sería dependiente del modelo usado y por lo tanto equivalente a una metodología de control de calidad de los datos cuyo resultado final sería una muestra de datos históricos corregidos con un modelo en particular, en vez de un modelo idóneo seleccionado a partir de la comparación de sus bondades con las de otros modelos.

La comparación entre estos dos diagnósticos nos lleva a la pregunta ¿qué sucedería si corregimos los datos diagnosticados como *outliers* originados por problemas en los datos cuando son *outliers* originados por problemas de especificación en el modelo? La respuesta favorece la validez de un diagnóstico intermedio que use una definición general de *outliers*, como observaciones fuera de patrón previamente establecido por el modelo y resultado de choques o errores fuera de cierto intervalo. La solución es reducir la importancia de su influencia adversa sobre los estimados mediante el uso de estimadores robustos o mediante la elección de una transformación de los datos originales que permita diluir estos choques idiosincráticos para capturar únicamente los cambios en la tendencia local³.

Este diagnóstico intermedio hace evidente que el esfuerzo en obtener proyecciones robustas, mediante estimadores robustos o procedimientos robustos en general, debe evitar el aumento en la precisión de las proyecciones de manera artificial, es decir, como resultado de "correcciones" en los datos históricos basados en un modelo individual. Se deduce que toda evaluación *ex post* de las proyecciones robustas debe realizar el cálculo de las RECM's a partir de los datos originales. Los procedimientos

³ Los resultados del presente trabajo enfatizan el uso de estimadores robustos para los modelos de predicción desagregada de la inflación IPC. aunque estrategias de transformación de las observaciones originales pueden ser particularmente útiles para el caso del crecimiento del Producto Bruto Interno (PBI). así como para el de la inflación del Índice de Precios al por Mayor (IPM).

robustos aplicados en el presente trabajo se apoyan en este diagnóstico intermedio que evita cualquier modificación de los datos históricos en la etapa de evaluación *ex post*.

En la primera sección se presenta una revisión de los estimadores robustos para modelos estáticos uni-ecuacionales y multi-ecuacionales. Esta revisión es el punto de partida para el diseño de procedimientos robustos en la estimación de modelos Sparse VAR, previéndose incluir otros estimadores robustos para estos modelos dinámicos cuando resulten útiles y baratos en términos del costo computacional. En la segunda sección se introducen las modificaciones básicas necesarias para aplicar algunos procedimientos robustos descritos en la primera parte para el caso de los modelos SparseVAR. En la tercera sección se realiza una comparación de la sensibilidad de la precisión *ex post* y de la precisión *ex post* para las proyecciones de cuatro modelos Sparse VAR, aplicados a la predicción de agregados como la inflación IPC con datos desagregados. Finalmente, se brinda un balance de las perspectivas a partir de estos resultados.

1. ESTIMADORES ROBUSTOS EN MODELOS ESTÁTICOS

1.1. UNI-ECUACIONALES

Los estimadores robustos aplicados a modelos uni-ecuacionales partieron de los estimadores robustos para la *media* (parámetro de posición) y la *desviación estándar* (parámetro de escala), originalmente propuestos para observaciones independientes e idénticamente distribuidas $x_i : F$, donde F es una clase de distribución simétrica cualquiera. Así, una de las definiciones originales de robustez corresponde a la ventaja de un estimador en desempeñarse mejor que la *media muestral* bajo muestras provenientes de una clase de distribuciones simétricas F con una masa en las colas mayor que en las de una distribución normal estándar, es decir, con una mayor probabilidad asignada a las observaciones extremas (observaciones influyentes por situarse “en las colas”).

Un ejemplo típico de estimador robusto en el contexto del parámetro de posición es la **mediana**, la cual presenta varianzas asintóticas menores que las de la media muestral cuando las observaciones provienen de una distribución Laplace o Cauchy⁴. Esta menor dispersión (mayor eficiencia) para distribuciones con colas anchas es una característica inherente a los estimadores robustos, ya que por construcción resultan menos sensibles ante la presencia de observaciones extremas.

De este modo, los estimadores robustos proveen de un seguro contra datos extremos (que no provendrían de la distribución normal estándar, con muy baja probabilidad) y a la vez se mantiene cerca del verdadero valor de la media poblacional cuando los datos provienen de una distribución normal estándar. Un ejemplo común en la literatura es tomado de Huber (1981) en el contexto del parámetro de escala: si los datos provienen de $N(\mu, \sigma^2)$, $d_n = \sum_{i=1}^T |x_i - \bar{x}| / n$ es un estimador de la dispersión que es casi 89 por ciento tan eficiente como el estimador $s_n = \sqrt{\sum_{i=1}^T ((x_i - \bar{x})^2) / n}$; pero si

⁴ Ver Amemiya (1985), quien enfatiza que la definición original de robustez requiere hacer referencia a la clase específica de distribuciones simétricas bajo consideración, debe tenerse en consideración cuando reseñemos los dos primeros grupos de estimadores robustos según Huber(1981).

sólo 0.2 por ciento de los datos provienen de $F = N(\mu, 9\sigma^2)$, entonces el estimador d_n es más eficiente. Otra referencia importante es Hampel (1973), quien afirma que los estimadores robustos brindan "un resguardo contra errores inesperadamente grandes, limitando la influencia de contaminaciones escondidas y observaciones dudosas o problemáticas, separando los *outliers* evidentes para un tratamiento separado (si se desea) y sin embargo manteniéndose cerca de la optimalidad para el modelo paramétrico en consideración".

Además de la mediana, la **media α -acotada**, que excluye las observaciones extremas en un porcentaje α del total de observaciones, constituye toda una familia de estimadores robustos en el contexto del parámetro de posición. Sin embargo, la media α -acotada y la mediana son sólo dos (2) de un conjunto muy amplio de estimadores robustos, de entre los cuales existe un estudio Monte Carlo para un total de sesenta y ocho (68) estimadores robustos.

Huber(1981) propone una taxonomía para este conjunto tan amplio de estimadores robustos en el contexto del parámetro de posición, clasificándolos en 3 grupos: los estimadores "tipo M" (seudo máxima verosimilitud), los "tipo L" (combinaciones lineales) y los "tipo R" (obtenidos de *rankings*). Resultará útil reseñar aquí sólo los dos primeros grupos de estimadores robustos, para los fines del presente estudio.

Grupo M

Consideremos una muestra $y_1, y_2, y_3, \dots, y_T$ de observaciones independientes e idénticamente distribuidas como F y con función de densidad f. Todo estimador en el grupo M se define como el escalar \hat{b}

$$\hat{b} = \arg \min_b \left\{ \sum_{t=1}^T \rho \left(\frac{y_t - b}{s} \right) \right\}$$

Si ρ es una función seleccionada tal que es doblemente diferenciable, su segunda derivada es continua por piezas y cumple con la condición $E \left\{ \rho' \left(\frac{y_t - \mu}{s_0} \right) \right\} = 0$ ⁵, es

posible usar una versión apropiada del teorema de la distribución asintótica de una función no lineal (donde ρ no dependa de T) para probar que el estimador b es asintóticamente normal con media μ y varianza asintótica

$$Asym \text{ var}(b) = \frac{1}{T} s_0^2 \frac{E \left\{ \left[\rho' \left((y_t - \mu) / s_0 \right) \right]^2 \right\}}{\left[E \left\{ \rho'' \left((y_t - \mu) / s_0 \right) \right\} \right]^2}$$

Usando el enfoque de decisión minimax, Huber(1981) eligió aquella función ρ que brindaba el mejor seguro contra la distribución menos favorable definida como cualquier miembro de la familia de distribuciones

$$F(z, \varepsilon) = (1 - \varepsilon)\Phi(z) + \varepsilon H(z) \quad z \equiv (y - \mu) / s$$

⁵ s es un estimado del parámetro de escala σ , mientras que $s_0 \equiv p \lim \{s\}$. Huber también propuso la alternativa de obtener b y s simultáneamente usando las condiciones de momentos muestrales asociadas a esta condición de momentos poblacional y a la correspondiente a la esperanza poblacional en el numerador del ratio en la fórmula de la varianza asintótica de b (calculada utilizando una función normal estándar).

donde $\varepsilon \in [0,1]$ es una constante, $\Phi(\cdot)$ es la distribución normal estándar y $H(\cdot) \in \Theta$, es decir, H es cualquier miembro de la familia de distribuciones simétricas Θ . La solución resultante fue

$$\rho(z) = \begin{cases} (1/2)z^2 & |z| < c \\ c|z| - (1/2)c^2 & |z| \geq c \end{cases}$$

donde c es una función de ε y puede usarse cualquier estimador robusto para el parámetro escala⁶.

Posteriormente, Andrews(1974) propuso la función

$$\rho(z) = \begin{cases} -(1 + \cos(z)) & |z| \leq \pi \\ 0 & |z| > \pi \end{cases}$$

y nuevamente puede usarse cualquier estimador robusto para el parámetro escala⁷.

Usando cualquiera de estas funciones, la generalización del estimador robusto M para un modelo de regresión lineal consiste simplemente en sustituir el escalar b por el escalar $x_i' b$ y definirlo como el vector $\hat{\beta}$

$$\hat{\beta} = \arg \min_b \left\{ \sum_{t=1}^T \rho \left(\frac{y_t - x_t' b}{s} \right) \right\}$$

donde el parámetro escala en la función ρ puede ser un estimador robusto del parámetro de escala σ como por ejemplo la versión truncada

$$s \equiv (2.1) \text{mediana} \left\{ |y_{(j)} - x_j' \hat{\beta}| : j \geq T - K \right\}$$

que elimina las menores K desviaciones debido a los K grados de libertad perdidos por la estimación de $x_j' \hat{\beta}$, donde $\hat{\beta} = \arg \min_b \left\{ \sum_{t=1}^T |y_t - x_t' b| \right\}$.

Grupo L

Todo estimador b en el grupo L se define como cualquier combinación lineal de los estadísticos de orden para la muestra $y_1, y_2, y_3, \dots, y_T$, definidos como

$$y_{(1)} \leq y_{(2)} \leq y_{(3)} \leq \dots \leq y_{(T)}$$

Así como todos los demás cuantiles muestrales, el cuantil 50 por ciento, o mediana, pertenece a este grupo ya que se define como

$$\text{mediana}(y_i) = \begin{cases} y_{\lfloor (T+1)/2 \rfloor} & T \text{ impar} \\ \text{convex}(y_{(T/2)}, y_{\lfloor (T/2)+1 \rfloor}) & T \text{ par} \end{cases}$$

donde $\text{convex}(a,b)$ es cualquier combinación convexa entre los puntos a y b. Asimismo, la media α -acotada también es una combinación convexa de los

⁶ Nótese que cuando $\rho(\cdot) = -\log(f(\cdot))$ y se utiliza la propuesta alternativa de Huber (ver nota anterior), se obtiene el estimador de máxima verosimilitud del parámetro de posición y escala. Por esta razón los estimadores robustos del grupo M también son conocidos como "tipo máxima verosimilitud". Asimismo, cuando $\rho(z) = z^2$, la fórmula de la varianza asintótica resulta ser la conocida varianza de la media muestral.

⁷ Dos estimadores robustos del parámetro de escala σ comúnmente usados en los estudios originales son $s \equiv (2.1) \text{mediana} \left\{ |y_i - \text{mediana}\{y_i\}| \right\}$ o una versión truncada que elimina la menor desviación debido a un grado de libertad perdido por la estimación de la mediana $\{y_i\}$.

estadísticos de orden pues asigna ponderaciones no-nulas a un porcentaje $1-\alpha$ de los mismos.

Antes de realizar la simple generalización del estimador robusto L para un modelo de regresión lineal resulta conveniente notar que el cuantil muestral de 100θ por ciento, $0 < \theta < 1$, pertenece al grupo L pues

$$q(\theta) = \begin{matrix} y_{(k)} & \min\{k \in \mathbb{Y} : k > T\theta\} & T\theta \in \mathbb{I} \\ \text{convex}(y_{(T\theta)}, y_{(T\theta+1)}) & & T\theta \in \mathbb{Y} \end{matrix}$$

donde el caso particular $q(0.5)$ corresponde a la mediana. Una definición alternativa de $q(\theta)$ es

$$q(\theta) = \arg \min_b \left\{ \sum_{y_i \geq b} \theta |y_i - b| + \sum_{y_i < b} (1-\theta) |y_i - b| \right\}$$

Se deduce que la generalización del estimador robusto L para un modelo de regresión lineal resulta de sustituir el escalar b por el escalar $x_i'b$ y definirlo como el vector $\hat{\beta}(\theta)$

$$\hat{\beta}(\theta) = \arg \min_b \left\{ \sum_{y_i \geq x_i'b} \theta |y_i - x_i'b| + \sum_{y_i < x_i'b} (1-\theta) |y_i - x_i'b| \right\}$$

Es importante notar que cualquier estimador que sea una combinación convexa de un conjunto finito de cuantiles también pertenece a este grupo (por ejemplo, el mismo $\hat{\beta}(\theta)$). Por esta razón, el miembro general de este grupo también es denominado “estimador de cuantiles”. El estimador de cuantiles es de gran utilidad para capturar particularidades específicas de la distribución de los errores, aunque su costo en ciertos contextos puede ser alto en términos de tiempo de computadora.

Antes de presentar los estimadores robustos análogos para sistemas de ecuaciones, debemos definir resumidamente dos conceptos que han resultado determinantes en los estudios teóricos y aplicados sobre estimadores robustos:

- a. **Punto de quiebre** (Hampel 1971): es aquella fracción aproximada de *outliers* presentes en la muestra que pueden ser tolerados sin que el estimador se pueda “quebrar”, es decir, se aleje demasiado del verdadero valor.
- b. **Función de influencia** (Hampel 1973): es aquella función que cuantifica el efecto sobre los estimados de la contaminación en cada observación.

Hasta el momento sólo se ha mencionado algunos de los estimadores robustos para el parámetro de escala, específicamente aquellos que pueden obtenerse a partir de un estimado inicial del parámetro de posición, como puede ser la mediana. Sin embargo, este procedimiento robusto en dos etapas resulta apropiado sólo en el caso de los estimadores robustos de posición del grupo L, pues estos estimadores son independientes del estimador de escala. Ello no sucede para el caso de los estimadores robustos de escala del grupo M⁸.

1.2. MULTI-ECUACIONALES

⁸ Los resultados preliminares demostraron que este “natural” procedimiento robusto no es lo más adecuado en el caso de los estimadores M, para los cuales se requiere utilizar procedimientos robustos donde todos los parámetros de posición y de escala sean obtenidos simultáneamente. La descripción de estos procedimientos se realiza más adelante.

Grupo M

La extensión de los estimadores robustos mencionados para el caso de un sistema de ecuaciones es directa, requiriendo únicamente de modificaciones evidentes. La forma reducida de un modelo estático puede escribirse como

$$y_t = B'x_t + \varepsilon_t$$

para $t \in \Delta_1 \equiv \{1, 2, \dots, T\}$, donde $y_t \in \mathbb{R}^k$, $x_t \in \mathbb{R}^Q$ y $B \in \mathbb{R}^{Q \times k}$. De manera similar al grupo M de estimadores para una regresión lineal, ahora partimos de una distribución simétrica cualquiera para el vector de errores $\varepsilon_t(B) \equiv y_t - B'x_t$ [Kx1] de un sistema de ecuaciones.

El estimador robusto requiere generalizar el supuesto usual de normalidad de los errores para el caso de que su matriz de segundos momentos pueda no existir. Se asume que $\varepsilon_t(B)$ se distribuye independiente e idénticamente con centro cero y "matriz de dispersión" Σ . Esto significa que la densidad de $\varepsilon_t(B)$ tiene la forma

$$f_{\varepsilon_t}(u(B), \Sigma) = \frac{g(u'(B)\Sigma^{-1}u(B))}{(\det(\Sigma))^{1/2}}$$

donde Σ es definida positiva (y denominada también *scatter matrix*) y g es una función positiva tal que permite que $f_{\varepsilon_t}(u(B))$ cumpla con los axiomas de probabilidad. Si el segundo momento de $\varepsilon_t(B)$ existe, entonces Σ será proporcional a la matriz de covarianzas de $\varepsilon_t(B)$. Los estimadores del grupo M para un sistema de ecuaciones pueden definirse como

$$(\hat{\beta}, \hat{\theta}) = \arg \min_{(\beta, \theta)} \left\{ -\sum_{t \in \Delta_1} f_{\varepsilon_t}(u(B), \Sigma) \right\}$$

donde $\beta \equiv \text{vec}(B)$ y $\theta \equiv \text{vech}(\Sigma)$ (nótese la redefinición sobre θ respecto a la notación usada en el grupo L arriba).

Sin embargo, la estimación simultánea de los estimadores M (posición y escala) resulta extremadamente costosa no sólo en términos del costo del tiempo de procesamiento, sino principalmente porque a medida que las dimensiones del sistema (K,Q) aumentan, se ha encontrado menores ventajas en robustez. La solución de que cada ecuación se estime por separado es la adecuada en este contexto, aunque la alternativa de usar estimadores generalizados M o GM *estimators* resulta más costosa en tiempo de procesamiento⁹.

⁹ Para el caso de modelos de series temporales en general, los estimadores M brindan robustez contra *outliers* innovacionales pero no contra *outliers* aditivos, salvo que utilicen la distribución de Student. Ver Franses, H. Kloek, T. & Lucas, A. (1999), quienes citan el énfasis de Martín (1979, 1981) a favor de los estimadores GM (*generalized-M*). La idea de los estimadores GM es similar a la de los estimadores M en el sentido de que ambos grupos aplican el principio de pseudo-máxima verosimilitud. Los estimadores GM en la literatura reciente también utilizan distribuciones del tipo $f_{\varepsilon_t}(u(B))$, como por ejemplo la distribución t de *Student* con $\nu \in \mathbb{N}_{++}$ grados de libertad. La "generalización" consiste en aplicar la idea detrás de la función de influencia de Hampel *et al* (1986). a la función de verosimilitud, la que entonces contiene una sumatoria ponderada en los términos que dependen de los datos y de los parámetros desconocidos. El objetivo de estos ponderadores es asignar un menor peso a las observaciones que tengan valores de $\varepsilon_t(B)$ de inusual magnitud y/o regresores x_t con valores inusuales. Los ponderadores se obtienen separadamente mediante una secuencia de iteraciones porque son funciones de una forma cuadrática en la posición y la escala multivariada del vector de

Grupo L

La generalización del grupo de estimadores L al caso de un sistema de ecuaciones no es tan directa. Para ello se requiere una muestra de escalares a partir de la cual se obtengan los estadísticos de orden y nuestros objetos ahora son vectores. Para ello aplicamos una medida de distancia “normalizada” a cada vector de errores $\varepsilon_i(B)$ disponible,

$$d_i^2 \equiv \varepsilon_i'(B)\Sigma^{-1}\varepsilon_i(B)$$

y estas son las denominadas “distancias de Mahalanobis”. Entonces, cualquier combinación lineal de los estadísticos de orden para la muestra de distancias de Mahalanobis, definidos como

$$d_{(1:T)}^2 \leq d_{(2:T)}^2 \leq d_{(3:T)}^2 \leq \dots \leq d_{(T:T)}^2$$

pertenece a este grupo de estimadores L para un sistema de ecuaciones como el ya definido. Nótese que esta es una expresión equivalente para los estadísticos de orden.

En el presente trabajo usaremos el estimador media α -acotada de estas distancias, que es una combinación convexa de estas distancias al asignar ponderaciones no-nulas a un porcentaje $1-\alpha$ de las mismas (y por lo dicho anteriormente, también es un “estimador de cuantiles”). Específicamente usaremos el estimador propuesto por Joossens & Croux (2005) para modelos VAR, que denominan “estimador de mínimos cuadrados acotados multivariados” o *Multivariate Least Trimmed Squares* (MLTS), que fuera introducido como estimador de sistemas de ecuaciones por Agulló, Croux & Van Aelst (2002).

El estimador MLTS se define como el argumento que minimiza una suma acotada de distancias Mahalanobis cuadradas (alternativamente, una media acotada de las mismas). Específicamente, minimiza la suma de las h menores distancias de Mahalanobis

$$(\hat{\beta}, \hat{\theta}) = \arg \min_{\{(\beta, \theta): |\Sigma|=1\}} \left\{ \sum_{s=1}^h d_{(s:T)}^2(B, \Sigma) \right\}$$

Joossens & Croux (2005) argumentan que es un consenso en la literatura que los estimadores M tienen un problema con dimensiones K elevadas. Asimismo sostienen que el estimador MLTS es bastante robusto en el contexto de modelos VAR(K, L). Sin embargo, los primeros autores definen $x_t \equiv [1, y'_{t-1}, y'_{t-2}, \dots, y'_{t-L}]'$, $x_t \in \mathbb{R}^Q$ donde $Q \equiv LK + 1$, y ahora $t \notin \Delta_1$ sino $t \in \Delta_{L+1} \equiv \{L+1, L+2, \dots, T\}$, por lo que las formulas anteriores se modifican ligeramente.

Agulló, Croux & Van Aelst (2002) demostraron que el estimador MLTS puede obtenerse utilizando el algoritmo “rápido” basado en el estimador del Mínimo Determinante de Covarianza (MCD), que es un estimador robusto de la escala multivariada que se aplicará al vector de errores del sistema. Joossens & Croux (2005) definen este algoritmo como una búsqueda estocástica de las h

regresores estacionarios x_t . Las iteraciones son necesarias porque la posición y la escala vectoriales son actualizados como promedios ponderados de x_t con los ponderadores de la iteración anterior, partiendo de estimadores iniciales de la posición y la escala.

"observaciones" que minimizan el determinante de la matriz de covarianzas de los residuos correspondientes a dichas "observaciones, donde los residuos se obtienen con el estimador de MCO, típicamente usado para los parámetros de un sistema de ecuaciones como el VAR.

Formalmente, definamos $Z \equiv \{(x_t, y_t) : t \in \Delta_{L+1}\} \subset \mathbb{R}^{K+Q}$ como el conjunto de datos y $\Lambda(h) \equiv \{H \subset \Delta_{L+1} : \#H = h\}$ como la familia de todos los subconjuntos de Δ_{L+1} con el mismo número de elementos h. Para cada subconjunto $H \in \Lambda(h)$, podemos definir $\hat{B}_{MCO}(H)$ como el estimador MCO obtenido con las h observaciones del subconjunto H,

$$\hat{B}_{MCO}(H) \equiv [X'(H)X(H)]^{-1} X'(H)Y(H)$$

donde las matrices X(H) e Y(H) son las sub-matrices de X e Y que resultan de seleccionar las observaciones (filas) según el índice H, respectivamente. Definiendo $\varepsilon[\hat{B}_{MCO}(H)] \equiv Y(H) - \hat{B}'_{MCO}(H)X(H)$, el estimador correspondiente de la matriz *scatter* obtenido de este subconjunto de observaciones es

$$\hat{\Sigma}_{MCO}(H) \equiv \frac{1}{h-K} \varepsilon[\hat{B}_{MCO}(H)]' \varepsilon[\hat{B}_{MCO}(H)]$$

de donde el problema es hallar el índice H de observaciones que minimiza el determinante de este estimador,

$$\hat{H} \equiv \arg \min_{H \in \Lambda(h)} |\hat{\Sigma}_{MCO}(H)|$$

por lo tanto, según la demostración de equivalencia dada en Agulló, Croux & Van Aelst (2002), el estimador MLTS para un modelo VAR (sistema en forma reducida) puede definirse como:

$$\hat{\beta}_{MLTS}(Z) \equiv \text{vec}(\hat{B}_{MCO}(\hat{H}))$$

$$\hat{\theta}_{MLTS}(Z) \equiv \text{vech}(\hat{\Sigma}_{MCO}(\hat{H}))$$

Joossens & Croux (2005) argumentan que la eficiencia del estimador MLTS es bastante baja, por lo que recomienda usar una versión "re-ponderada" para mejorar su desempeño en muestras pequeñas, que es análoga a la "generalización" de los estimadores M sobre la base del concepto de la *función de influencia* de Hampel *et al* (1986), esta vez aplicada sobre las distancias cuadradas de Mahalanobis. Para obtener este estimador "reponderado" MLTS (ó estimador RMLTS) definimos

$$J \equiv \left\{ j : d_j^2(\hat{B}_{MLTS}, \hat{\Sigma}_{MLTS}) \leq q_\delta \right\}$$

donde $q_\delta \equiv \chi_{Q,1-\delta}^2$ es el cuantil superior δ -por ciento de una distribución χ_Q^2 . δ es simplemente la fracción de acotación de la etapa de reponderación. Entonces el estimador RMLTS se define como

$$\hat{\beta}_{RMLTS}^{MCO}(Z) \equiv \text{vec}(\hat{B}_{MCO}(J))$$

$$\hat{\theta}_{RMLTS}^{MCO}(Z) \equiv \text{vech}(c_\delta \hat{\Sigma}_{MCO}(J))$$

donde c_δ es un factor de consistencia que permite una estimación consistente de Σ en la distribución simétrica general para los errores, mencionada arriba, que será especificado más adelante en las aplicaciones.

La “reponderación” simplemente considera que la presencia de algunos *outliers* puede persistir luego de la primera etapa que obtiene el estimador MLTS. Si esto sucede, estos *outliers* persistentes serán detectados por sus elevados errores cuadráticos normalizados de Malahanobis, y por ende, las “observaciones” asociadas serán excluidas bajo la definición de las que deben permanecer en J y en el cómputo del estimador RMLTS.

2. ESTIMACIÓN DE MODELOS SPARSE VAR ROBUSTOS

La construcción de los modelos SparseVAR como los descritos en Barrera (2005, 2006) requiere básicamente de dos etapas. La primera etapa consiste en la especificación de restricciones de exclusión y la segunda etapa, en la estimación del sistema sobre la base de estas restricciones.

La primera etapa consiste en una búsqueda heurística de los rezagos de las variables en forma uni-ecuacional que minimice una versión del criterio de información de Akaike (AIC) modificado para muestras pequeñas. En el contexto de estimaciones robustas a la presencia de *outliers*, se utilizó el estimador M de Andrews(1974) así como el estimador MCD para la escala univariada de los errores estimados.

Asimismo, usar el criterio de AIC no es exacto en este contexto ya que no se parte de distribuciones normales; sin embargo, usar el estimador del parámetro escala en vez de la varianza estimada en la “función penalizada” de Akaike es una práctica común en el contexto de los estimadores M como el de Andrews. Por otra parte, la ligera modificación de usar el estimador MCD para la escala univariada de los errores estimados a partir de este estimador M es muy similar a la propuesta de Preminger & Sakata(2001) de usar el criterio de información PSC (“criterio de escala S penalizada” en el contexto de la familia de estimadores S¹⁰).

Obteniendo las restricciones de exclusión, la segunda etapa consiste en calcular el estimador RMLTS ya descrito para el caso de una estimación GLS estimada (EGLS) que posibilita el uso de dichas restricciones. Lutkepohl(1993) utiliza la siguiente representación de las restricciones en un sistema multivariado

$$\beta \equiv \text{vec}(B) = R\gamma + r$$

donde γ es un vector [Mx1] de M parámetros diferentes de cero, R es una matriz [K(KL+1)xM] conocida de restricciones y r es un vector [K(KL+1)xM] de coeficientes conocidos (ceros en nuestro caso). Si conocemos la matriz de covarianzas de los errores del sistema, Σ_u , el estimador GLS se define como

$$\hat{\gamma} \equiv \arg \min_{\beta} \left\{ u'(\beta) (I_T \otimes \Sigma_u^{-1}) u(\beta) \right\}$$

es decir,

$$\hat{\gamma}_{GLS} = \left[R' (XX' \otimes \Sigma_u^{-1}) R \right]^{-1} R' (XX' \otimes \Sigma_u^{-1}) y \quad y \equiv \text{vec}(Y)$$

¹⁰ La familia de estimadores S minimiza la escala S muestral de los errores estimados de manera análoga a la minimización del error cuadrático medio contemporáneo (ECM(h=1)) de los errores estimados mediante el estimador MCO. La propuesta de Preminger & Sakata(2001) es sustituir la raíz del ECM(1) en la función penalizada de Akaike por el estimado de la escala S. Para completar la explicación, debemos mencionar que los estimadores que utilizan la función de influencia de Hampel pertenecen a esta familia y que el estimador M de Andrews es equivalente a una primera iteración de un estimador S. Ver Amemiya(1985).

Sin embargo, la matriz de covarianzas es desconocida. Si utilizamos un procedimiento iterativo que parte de una matriz idéntica, se usa la fórmula anterior para hallar los errores estimados y luego se utiliza un estimador consistente de la matriz de covarianzas de los mismos, $\hat{\Sigma}_u$, el estimador EGLS resulta

$$\hat{\gamma}_{EGLS} = \left[R' \left(XX' \otimes \hat{\Sigma}_u^{-1} \right) R \right]^{-1} R' \left(XX' \otimes \hat{\Sigma}_u^{-1} \right) y$$

entonces nuestro estimador RMLTS-EGLS resulta ser

$$\hat{\gamma}_{RMLTS}^{EGLS} (Z) \equiv \text{vec} \left(R \hat{\gamma}_{EGLS} (J) \right)$$

$$\hat{\theta}_{RMLTS}^{EGLS} (Z) \equiv \text{vech} \left(c_{\delta} \hat{\Sigma}_{EGLS} (J) \right)$$

donde, como podrá ser evidente,

$$\hat{\gamma}_{EGLS} (H) = \left[R' \left(X(H)X'(H) \otimes \hat{\Sigma}_u^{-1}(H) \right) R \right]^{-1} R' \left(X(H)X'(H) \otimes \hat{\Sigma}_u^{-1}(H) \right) y(H)$$

Finalmente, como todo estimador RMLTS, se requiere definir el valor de α , es decir, el porcentaje de *outliers* contra los cuales el estimador RMLTS tiene resistencia sin que se "quiebre". Una primera estrategia para determinarlo es usar el número mínimo de "outliers" que *a priori* se puede tener en la muestra. En nuestro caso, considerando 6 para 1998 y 6 para 2006, se obtiene 8.63 por ciento (12 entre 139 observaciones).

Sin embargo, también se busca minimizar el número de muestreos aleatorios del algoritmo de optimización estocástica para ahorrar el costo en tiempo de computadora. Si el número de *outliers* que *a priori* se piensa que existe en la muestra es alto, la búsqueda estocástica requerirá un mayor número de iteraciones para elevar la probabilidad de que se localice grupalmente dicho mayor número de *outliers*. Por ello se establece una estrategia adaptable que consiste en usar un bajo porcentaje α "inicial", de 3 por ciento, debido a que se ha utilizado el conjunto de todos los parámetros individualmente estimados (ecuación por ecuación) como un adecuado estimado "de partida" para la estimación RMLTS-EGLS. De esta manera, existe una mayor confianza de que los *outliers* remanentes en la primera etapa (definida por α) serán eliminados de todas maneras en la segunda etapa por el algoritmo interno de "re-ponderación". Además del costo en tiempo de computadora, la razón para un bajo α inicial es que así los estimados RMLTS-EGLS lograrán una elevada eficiencia para muestras pequeñas (de hecho, la eficiencia de los estimadores EGLS también es mayor que la de los estimadores MCO bajo condiciones de normalidad).

Además siempre será posible adaptar el nivel de α , el 'seguro contra *outliers*', posteriormente, según se identifique mejor qué modelos específicos pueden adolecer de un mayor número de *outliers*.

3. SENSIBILIDAD Y PRECISIÓN COMPARADAS

3.1. SENSIBILIDAD EX POST

En la introducción se presentó la sensibilidad de la precisión de las proyecciones para cuatro horizontes de proyección $h \in \{1, 3, 6, 9\}$, correspondientes a 4 modelos Sparse VAR "plain vainilla" ante la presencia de una secuencia de *outliers* recientemente experimentada. Dichas secuencias de los RECM(h) se construyeron con el mismo

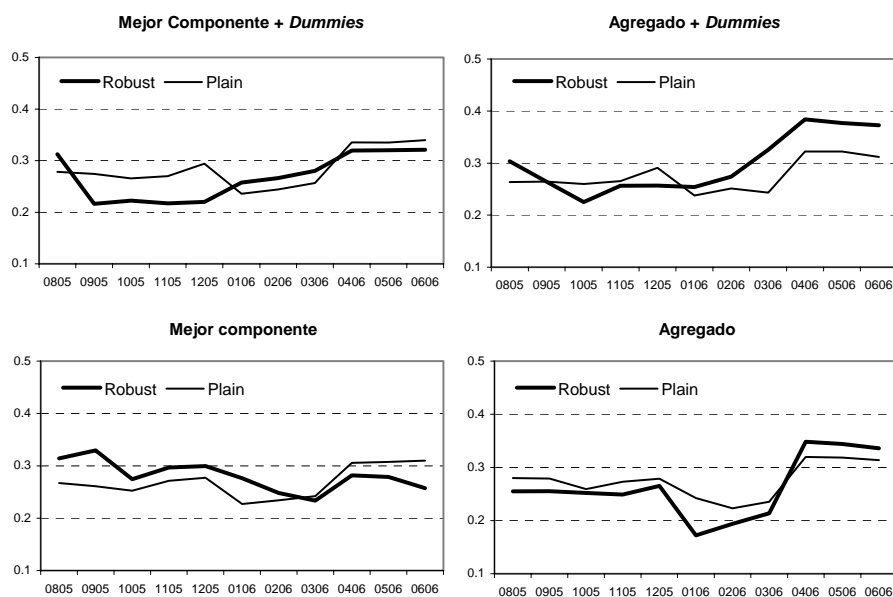
número de errores de predicción ejecutados para el horizonte especificado, por lo que sus cambios sólo reflejan el impacto de la secuencia de *outliers* a medida que aparecen en la muestra.

En esta sección comparamos esas secuencias con las correspondientes a los modelos Sparse VAR robustos obtenidos a partir de estimadores RMLTS-EGLS, ya descritos. Resultará evidente que las particularidades inherentes a estos cuatro modelos del SPD se reflejarán en sensibilidades potenciales diferentes (con o sin *dummies* estacionales, con la elección del mejor componente asociado para cada ecuación del VAR o simplemente el uso del agregado como el componente asociado).

El Gráfico 2 presenta estas comparaciones para las secuencias de RECM($h=1$). Con un horizonte $h=1$, los dos modelos robustos que eligen el mejor componente en cada ecuación (primera columna) presentan secuencias RECM muy cercanas a los correspondientes modelos “plain”. El que no incluye *dummies* estacionales es inicialmente sensible a la secuencia de *outliers* para luego mejorar en algo su precisión respecto al correspondiente modelo “plain”. Por su parte, los dos modelos robustos que eligen el agregado en cada ecuación (segunda columna) presentan una mayor precisión respecto a los correspondientes modelos “plain”, que luego se deteriora. Sólo en el caso del que no incluye *dummies* estacionales, esta ventaja se deteriora a partir de abril 2006.

Estos resultados demuestran que, para el porcentaje fijo de acotación de las distancias de Mahalanobis de $\alpha = 0.03$, no existe gran diferencia en la sensibilidad de las proyecciones de inflación IPC en horizontes de muy corto plazo. Una manera de reducir la sensibilidad de RECM($h=1$) a medida que aparecen los *outliers* puede ser aumentar este porcentaje a $\alpha = 0.06$, asumiendo el costo asociado en términos de eficiencia.

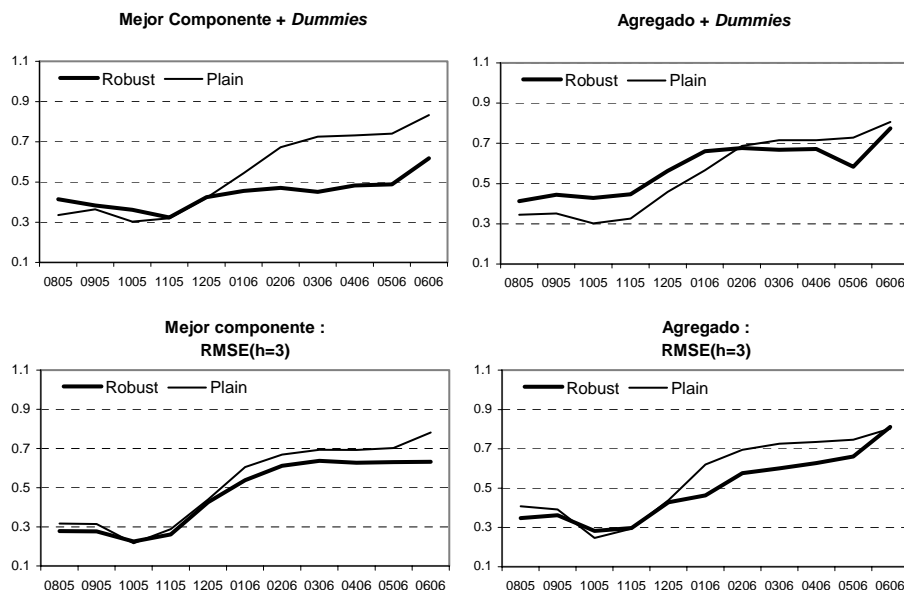
Gráfico 2
Inflación IPC : RMSFE($h=1$)



El Gráfico 3 realiza estas mismas comparaciones para los RECM($h=3$). Con un horizonte $h=3$, los dos modelos robustos que eligen el mejor componente en cada ecuación (primera columna) presentan secuencias RECM por debajo de los correspondientes modelos “plain” a medida que se presenta la secuencia de *outliers*.

Sólo el que incluye *dummies* estacionales mejora sustancialmente su precisión aunque adolezca de cierta sensibilidad. De los dos modelos robustos que eligen el agregado en cada ecuación (segunda columna), sólo el que no tiene *dummies* presenta una mayor precisión durante casi toda la parte de la secuencia, respecto a su contraparte “plain”.

Gráfico 3
Inflación IPC : RMSFE(h=3)



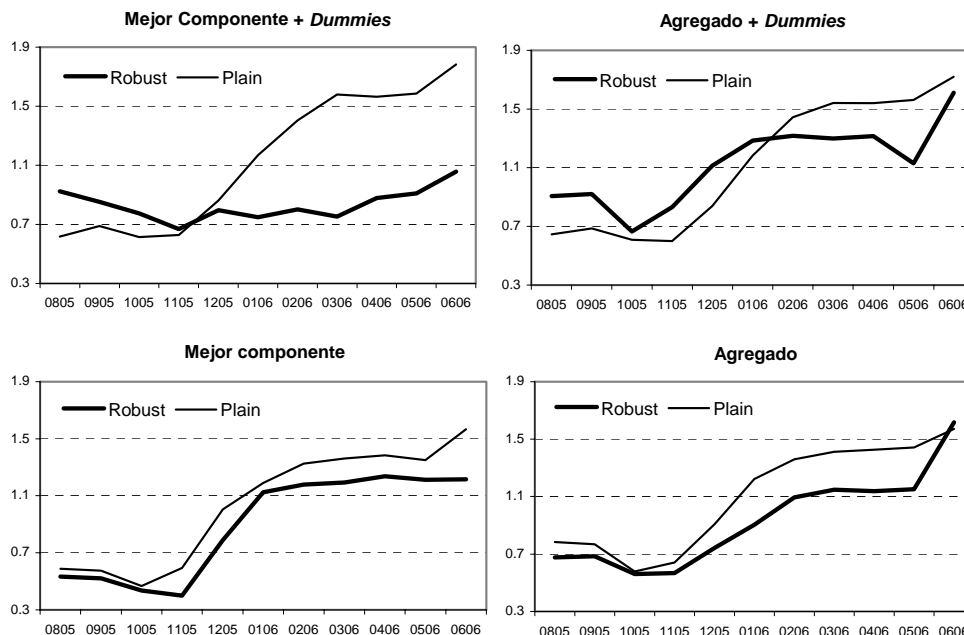
Estos resultados demuestran que, para el porcentaje fijo de acotación de las distancias de Mahalanobis de $\alpha = 0.03$, se logra una reducción importante en la sensibilidad de las proyecciones de inflación IPC con un horizonte ligeramente más amplio. Es posible reducir la sensibilidad de RECM(h=3) a medida que aparecen los *outliers* si se usa un porcentaje de acotación de $\alpha = 0.06$, asumiendo el costo asociado en términos de eficiencia.

Estas conclusiones se mantienen a medida que el horizonte de proyección aumenta. Específicamente, la precisión de los modelos robustos resulta mayor y aunque persiste cierta sensibilidad ante la secuencia de *outliers*, ésta se reduce con el horizonte. La persistencia de cierto grado de sensibilidad obtenida para horizontes relativamente cortos refleja que el estimador RMLTS-EGLS requiere aumentar el porcentaje de acotación aplicada a las distancias de Mahalanobis, lo que también el hecho de que dichos *outliers* no necesariamente impactan rápidamente sobre los modelos “plain”. Un aumento en este porcentaje puede favorecer el desempeño de aquellos modelos robustos que aún no presentan mejoras sustanciales respecto a sus contrapartes “plain”.

Por otro lado, en la introducción se argumentó que las observaciones fuera del patrón previamente establecido tuvieron lugar en mayores expectativas de inflación en el mediano plazo, principalmente, lo que, sin embargo, impactó en la formación de los precios de bienes incluidos en la inflación subyacente IPC y, en mayor medida, en la inflación no subyacente IPC. También se mencionó que hacia el final de la secuencia de *outliers*, el deterioro en la precisión de las proyecciones “plain” del agregado inflación IPC se debe a la menor precisión de las proyecciones de los componentes de la inflación no subyacente IPC. Para obtener una explicación de la mejora establecida en la sensibilidad de las proyecciones de inflación IPC, el Gráfico 4

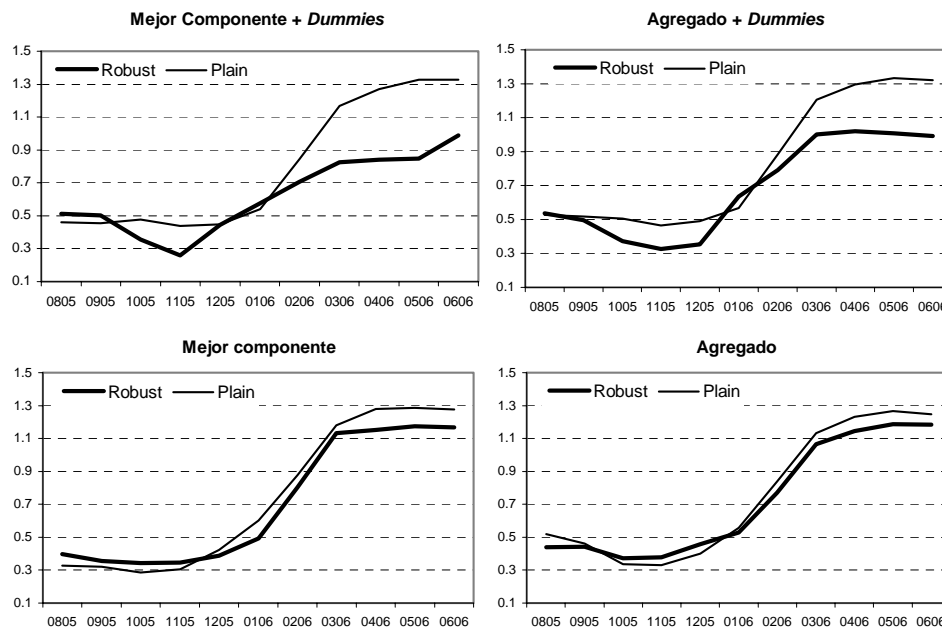
compara las secuencias de RECM(h=3) para el sub-agregado de inflación no subyacente IPC.

Gráfico 4
Inflación No Subyacente IPC : RMSFE(h=3)



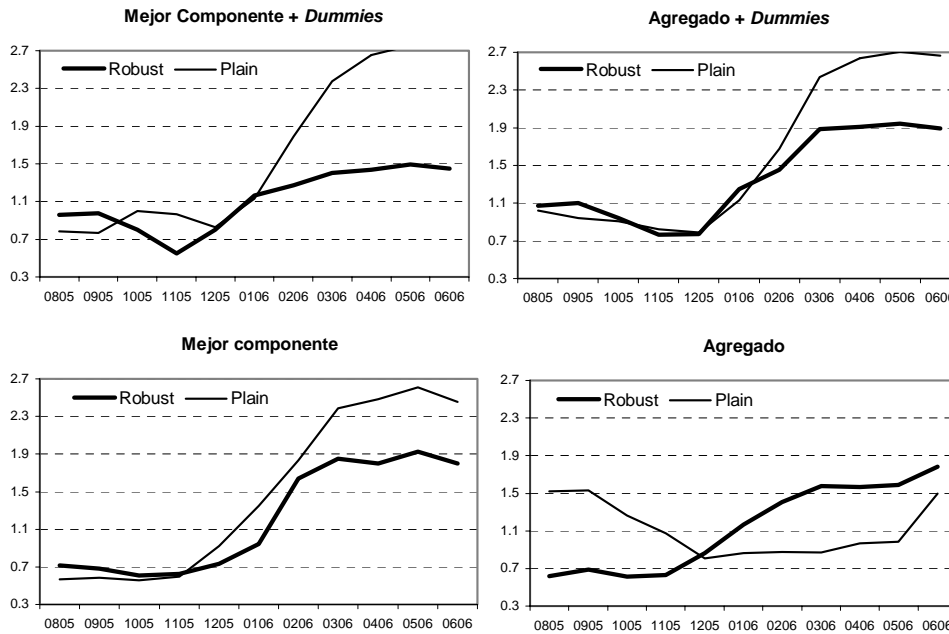
Este gráfico indica que la ventaja de las proyecciones robustas para el agregado de Inflación IPC se explica justamente por las mejoras importantes en la sensibilidad de las proyecciones robustas de la inflación no subyacente, especialmente en el tipo de modelo que incluye *dummies* estacionales y selecciona el mejor componente asociado para cada ecuación del VAR.

Gráfico 5
Inflación IPC : RMSFE(h=6)



Algunos de estos resultados se mantienen en el caso de los otros dos horizontes $h \in \{6,9\}$, aunque todavía resulta ser tentativo para el último caso (dado el reducido número de errores ejecutados utilizado para calcular la secuencia). Es importante notar que para el horizonte $h=6$, los RECM($h=6$) de los modelos robustos también demuestran una sensibilidad excesiva ante la secuencia de *outliers*, lo que confirma que se deberá elevar el porcentaje de acotación a $\alpha = 0.06$ ó $\alpha = 0.09$, a pesar del costo de la menor eficiencia. El origen de esta sensibilidad resulta ser un exceso de *outliers* en la inflación no subyacente.

Gráfico 6
Inflación No Subyacente IPC : RMSFE($h=6$)



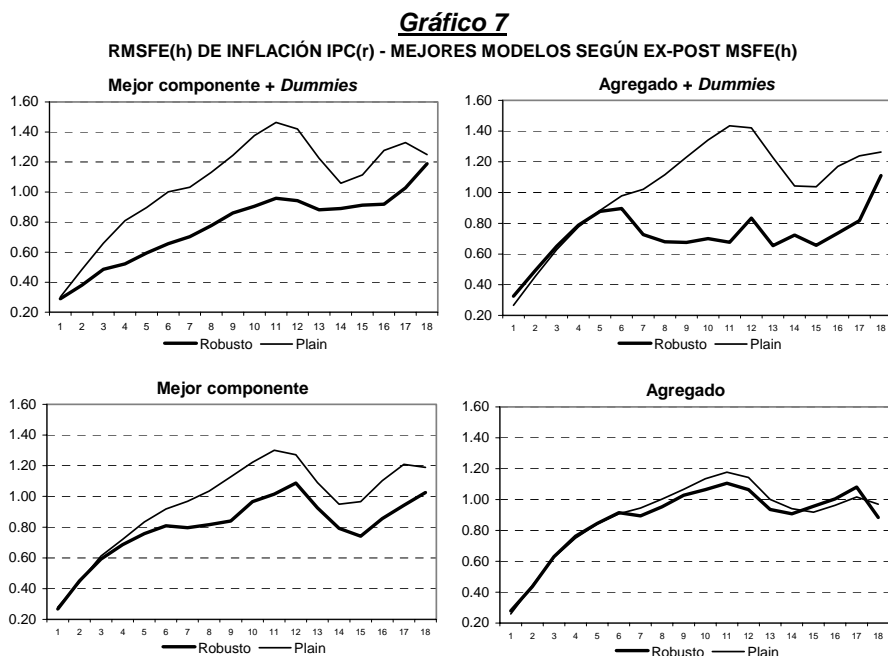
En esta sección se comparó la sensibilidad de la precisión de las proyecciones de los modelos “plain” y los modelos robustos para tres horizontes de corto plazo. En la medida que la sensibilidad de mediano plazo, digamos $h=9$, por lo general resulta relativamente proporcional en horizontes cercanos pero más amplios, los resultados obtenidos pueden ser extensivos *grosso modo* a la sensibilidad de mediano plazo. Aún no se dispone de un número suficiente de errores de predicción ejecutados para horizontes $h=12$ como para construir gráficos similares a los anteriores y afirmarlo categóricamente, aunque los resultados mostrados confirman la necesidad de ajustar el porcentaje de acotación, además de las ventajas de los estimadores robustos¹¹.

¹¹ El estimador robusto utilizado, una variante de los MCD, pertenece a un conjunto de estimadores robustos que son computacionalmente difíciles de calcular (“NP-hard” según Bernholt(2006)). Este costo debe tenerse siempre presente al determinar los parámetros de diseño de los cuales depende, como el porcentaje de acotación. Es posible que la baja probabilidad de obtener cálculos correctos para ciertas muestras de datos refleje la necesidad de modificar estos parámetros de diseño. Afortunadamente, los resultados presentados proporcionan una prueba de que, para la muestra de datos utilizada, se ha logrado encontrar modelos correctos que proporcionan un seguro de robustez respecto a la presencia de *outliers* con una “prima” relativamente barata. La elevación del porcentaje de acotación sugerida puede constituir un análisis de sensibilidad importante para sugerir alternativas de cómputo viables frente a las dificultades para obtener cálculos correctos en un tiempo polinómico no determinístico.

Sin embargo, con los errores de predicción ejecutados disponibles es posible comparar el desempeño en predicción de ambos tipos de modelos, teniendo en cuenta que el número de errores ejecutados usado en cada estadístico RECM(h) decrece a medida que aumenta el horizonte.

3.2. PRECISIÓN EX POST

La comparación de la precisión *ex post* de los diferentes modelos para diferentes horizontes de proyección es útil para deducir qué modelo predice mejor en qué horizonte. Esta evaluación considera los errores de predicción ejecutados más recientes sin ajustar los datos por la presencia de *outliers*, lo que es consistente con la definición utilizada en el presente trabajo. Los resultados se presentan en el Gráfico 7.



Los dos modelos robustos que eligen el mejor componente en cada ecuación (primera columna) presentan RECM(h) consistentemente menores a los de los correspondientes modelos “plain”. Nótese que el modelo que incluye *dummies* estacionales mejora sustancialmente su precisión, a pesar de cierto grado de sensibilidad a la presencia de *outliers* para horizontes cortos. Ello se debe básicamente al adecuado desempeño de sus proyecciones internas de inflación no subyacente IPC, como puede verificarse en los gráficos siguientes.

De los dos modelos robustos que eligen el agregado en cada ecuación (segunda columna), sólo el que incluye *dummies* estacionales es más preciso que su contraparte “plain” para todos los horizontes. Este modelo presenta un desempeño muy similar al que elige el mejor componente en cada ecuación (primera columna, primera fila), excepto en horizontes cortos, debido al adecuado desempeño de sus proyecciones internas de inflación subyacente IPC hasta $h=10$, lo que fue compensado con el desempeño intermedio de sus proyecciones internas de inflación no subyacente IPC para los horizontes cortos (ver Gráficos 8 y 9).

Gráfico 8

RMSFE(h) DE INFLACIÓN IPC(r) SUBYACENTE - MEJORES MODELOS SEGÚN EX-POST MSFE(h)

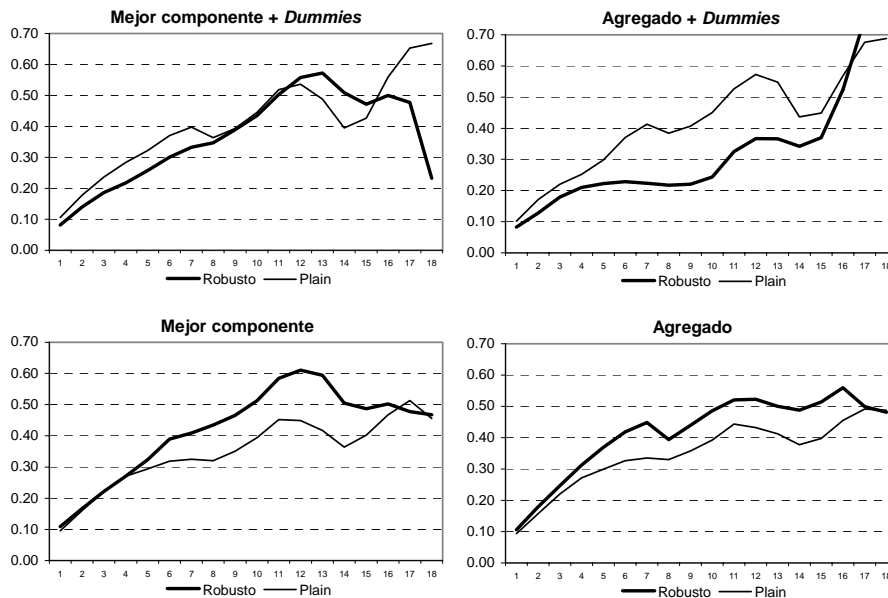
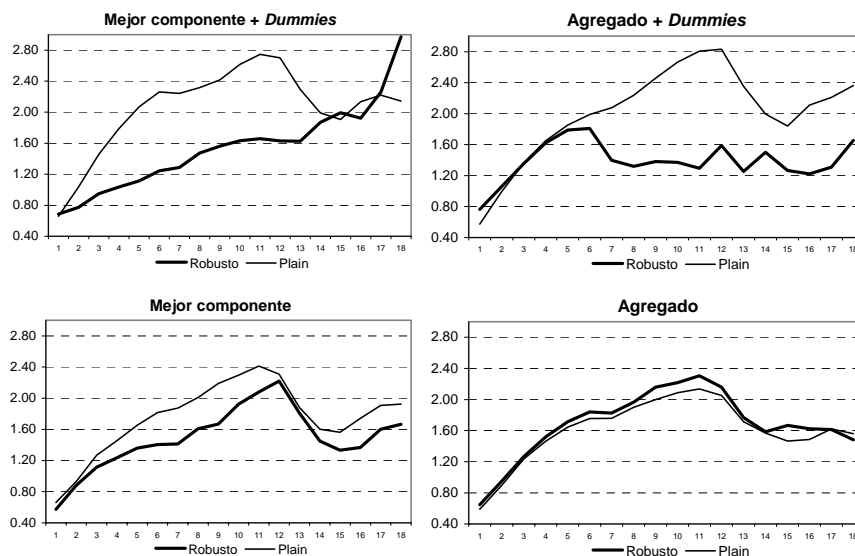


Gráfico 9

RMSFE(h) DE INFLACIÓN IPC(r) NO SUBYACENTE - MEJORES MODELOS SEGÚN EX-POST MSFE(h)



5. CONCLUSIONES

Las proyecciones robustas, obtenidas de modelos Sparse VAR robustos, tienen una menor sensibilidad ante secuencias de *outliers* como las recientes o como las experimentadas a inicios de 1998 durante el último fenómeno de “El Niño”. Por lo tanto, se está mejor preparado para recibir otro a inicios del 2007 sin afectar las proyecciones sobremanera.

Sin embargo, queda pendiente comparar estos resultados con los obtenidos luego de aumentar el porcentaje de acotación del estimador RMLTS-EGLS, lo que puede mejorar el desempeño de varios modelos, tanto en lo referido al grado remanente de sensibilidad de la precisión de las proyecciones para determinados horizontes como en la precisión *ex post* para los distintos horizontes. Los resultados obtenidos

buscaron minimizar el costo en términos de tiempo de computadora, ya que un porcentaje de acotación mayor requiere un número mayor de iteraciones en la búsqueda estocástica detrás del RMLTS-EGLS además de una pérdida en la eficiencia de estos estimadores.

Es importante señalar que el estimador RMLTS-EGLS utilizado recoge gran cantidad de información de la distribución muestral de los errores del sistema de ecuaciones correspondiente. Para probarlo debemos mencionar que, por ejemplo, el estimador RMLTS-EGLS es equivalente a una combinación convexa que otorga el mismo peso positivo a todos los cuantiles específicamente definidos como

$$q(\hat{\theta}) = d_{(\hat{k}:T)}^2 \min \left\{ k \in \mathbb{Y} \wedge T\hat{\theta} \in i : k > T\hat{\theta} \right\}$$

es decir, todos aquellos cuantiles que equivalen a las distancias de Mahalanobis con igual ponderación positiva en la media α -acotada detrás del estimador RMLTS-EGLS. Para ello se debe obtener primero este conjunto de cuantiles de modo tal que sean muy similares para distintos tamaños muestrales T . Por lo tanto, el estimador RMLTS-EGLS recoge la información de la distribución muestral que resulta de modificar el supuesto de normalidad por uno que captura un gran número de distribuciones empíricas con colas anchas, en el contexto de estimadores robustos¹².

Finalmente, tanto la evaluación de la sensibilidad como la evaluación de la precisión *ex post* se realizaron sin ajuste alguno sobre los datos. Esta estrategia permite seguir evaluando la secuencia de *outliers* en términos de un problema de especificación a resolverse con modelos de mayor dimensión (un conjunto de información más amplio). El beneficio de este camino intermedio de evitar ajustar los datos para luego hacer las evaluaciones es que se mantiene la comparabilidad de las evaluaciones a través de distintos modelos. El único costo es que los niveles de la precisión *ex ante* son sub-valorados ya que algunos errores de predicción ejecutados se miden con respecto a las observaciones fuera de patrón previamente establecido.

BIBLIOGRAFÍA

Agulló, José; Croux, Christophe & Van Aelst, Stefan (2002) The Multivariate Least Trimmed Squares Estimator. Working Paper.

Amemiya, Takeshi (1985) *Advanced Econometrics*, Harvard University Press.

Bernholt, Thorsten (2006) Robust Estimators are hard to compute, Working Paper, Lehrstuhl Informatik 2, Universität Dortmund.

Fox (1972) "Outliers in time series", *Journal of the Royal Statistical Society, Series B*, vol. 34(3), pgs. 350-363.

Franses, H. Kloek, T. & Lucas, A. (1999) Outlier robust analysis of long run marketing effects for weekly scanning data, en *Journal of Econometrics*, Vol. 89, pgs. 293-315.

Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., & Stahel, W.A. (1986) "Robust Statistics: The approach based on Influence Functions", Wiley, New York.

¹² Por supuesto que similar información de la distribución muestral será capturada al usar un estimador de cuantiles que no sea equivalente a una media α -acotada, pero ello tendría un costo computacional muy alto, como es típico para el estimador de cuantiles general definido en la sección 2.

Huber, P.J. (1981) Robust Statistics, John Wiley and Sons.

Joossens, Kristel & Croux, Christophe (2005) Robust Estimation of the Vector Autoregressive Model by a Trimmed Least Squares Procedure, Department of Applied Economics Working Paper, Universidad Católica de Lovaina, Mimeo.

Machado J.A.F. (1993). Robust model selection and M-estimation. *Econometric Theory*, 9, 478-493.

Preminger, Arie & Sakata, Shinichi (2005) A model selection method for S-estimation, Mimeo.