

UNA NOTA EMPÍRICA SOBRE *OUTLIERS* ADITIVOS Y NO ESTACIONARIEDAD EN SERIES DE INFLACIÓN EN AMÉRICA LATINA¹

Gabriel Rodríguez

1. Introducción

Cuando tests de raíz unitaria regulares o comunes, tales como el Dickey-Fuller aumentado (Dickey y Fuller 1979; Said y Dickey 1984) o el Phillips-Perron (Phillips 1987; Phillips y Perron 1988), se aplican a un conjunto de series de inflación de América Latina, se obtiene un fuerte rechazo de la hipótesis de raíz unitaria. Este resultado es invariable aun en el caso de utilizar un largo retardo, lo cual parece reflejar la existencia de distorsiones de nivel. Una inspección visual de las cuatro series de inflación latinoamericanas (véase la figura 1) indica la existencia de evidentes *outliers*. En este sentido, estas series ofrecen un excelente ejemplo de los efectos de grandes outliers sobre las propiedades de series temporales. Esto es aún más claro dado el hecho de que la mayor parte de la investigación que utiliza este conjunto de variables supone que dichas variables son $I(0)$. Esta nota ofrece un ejemplo de los peligros de emplear tests de raíz unitaria frecuentemente usados cuando existen grandes *outliers*.

Utilizando recientes métodos propuestos por Vogelsang (1999) y Perron y Rodríguez (2003), nuestros resultados muestran clara evidencia de la presencia de *outliers* aditivos en las cuatro series de inflación analizadas. La mayoría de esas

¹ Documento preparado en abril del 2006 en homenaje al profesor Máximo Vega-Centeno, profesor del Departamento de Economía de la Pontificia Universidad Católica del Perú. El autor fue alumno del profesor Vega-Centeno en el curso de Teoría del Crecimiento y asistente de docencia en el curso de Econometría 2. Asimismo, el profesor Vega-Centeno fue el supervisor de la tesis de Licenciatura del autor. El presente documento es una traducción del paper «An Empirical Note about Additive Outliers and Nonstationarity in Latin American Inflation Series», publicado por el autor en *Empirical Economics* 2004, 29, pp. 361-372. Mis agradecimientos a Pierre Perron por sus útiles comentarios a una versión preliminar del documento. Asimismo, los útiles comentarios de un anónimo árbitro, Baldev Raj, Yagadeesen Samy, son apreciados. Finalmente, agradezco a André Lucas por sus interesantes sugerencias relacionadas con el uso de su programa Robust Inference Plus Estimation (RIPE).

observaciones están asociadas a fechas de aplicación de programas de estabilización en anteriores años. Dichos programas fueron aplicados con el objetivo de parar los altos niveles de inflación en esos países. La aplicación del estadístico ADF corregido por la presencia de *outliers* aditivos confirma la estacionariedad de la tasa de inflación para los casos de Bolivia y Chile. Sin embargo, las series de inflación para Argentina y Perú muestran ser no estacionarias.

Es claro que existen muchos otros procedimientos para identificar la presencia de *outliers* aditivos. Por ejemplo, tenemos los métodos propuestos por Tsay (1986); Chang, Tiao y Chen (1988); Shin, Sharkar y Lee (1996); Chen y Liu (1993); y Gómez y Maravall (1992a, 1992b). Otra sugerencia interesante es aquella propuesta por Lucas (1995a, 1995b), y Hoek, Lucas y van Dijk (1995). Al final de este texto presentamos una breve comparación de nuestros resultados con aquellos obtenidos utilizando los métodos sugeridos por Lucas (1995a, 1995b).

Este documento está organizado de la siguiente manera. La sección 2 presenta el modelo, discute la detección de outliers y revisa brevemente los dos métodos propuestos por Perron y Rodríguez (2003) para detectar outliers aditivos. En la sección 3 se describe el estadístico ADF corregido por la presencia de outliers aditivos y se presentan los resultados del análisis empírico. La sección 4 presenta las conclusiones. Detalles sobre la fuente de la información son provistos en el apéndice.

2. La detección de *outliers* y tests de raíz unitaria con *outliers* aditivos

La detección de *outliers* en el marco teórico de una raíz unitaria es el contexto en el cual trabajan Vogelsang (1999) y Perron y Rodríguez (2003). El proceso generador de datos es de la siguiente forma:

$$y_t = d_t + \sum_{j=1}^m \delta_j D(T_{ao,j})_t + u_t \quad (1)$$

Donde $D(T_{ao,j})_t = 1$ si $t = T_{ao,j}$ y 0 si otro. Esto permite la presencia de m *outliers* aditivos que ocurren en las fechas $T_{ao,j}$ ($j = 1, \dots, m$). El término d_t especifica los componentes determinísticos. En la mayoría de los casos, $d_t = \mu$ si la serie no tiene tendencia o $d_t = \mu + \beta t$ si la serie tiene una tendencia. La función de ruido es integrada de orden uno; es decir, $u_t = u_{t-1} + v_t$, donde v_t es un proceso estacionario. El procedimiento de detección de *outliers*, sugerido por Vogelsang (1999), comienza con la estimación (por Mínimos Cuadrados Ordinarios-MCO) de la siguiente regresión (si es necesario, una tendencia lineal puede ser incluida),

$$y_t = \hat{\mu} + \hat{\delta} D(T_{ao})_t + \hat{u}_t \quad (2)$$

donde $D(T_{ao})_t = 1$ si $t = T_{ao}$ y 0 si otro. Asumamos que $t_{\delta}(T_{ao})$ denote el t-estadístico para evaluar $\delta = 0$ en (2). Siguiendo la recomendación de Chen y Liu (1993), la presencia de un *outlier* aditivo puede ser verificada utilizando $\tau = \sup_{T_{ao}} |t_{\delta}(T_{ao})|$.

Asumiendo que $\lambda = T_{ao} / T$ permanece fijo cuando T aumenta, Vogelsang (1999)

ha mostrado que cuando $T \rightarrow \infty$, la distribución límite de $t_{\delta}(T_{ao})$ es no estándar.² En Perron y Rodríguez (2003), valores críticos asintóticos para τ han sido obtenidos por simulaciones. Cuando existe un intercepto en (2), los valores críticos son 3,53, 3,11 y 2,92 al 1, 5 y 10% de significancia, respectivamente. Si una tendencia lineal es también incluida en (2), los valores críticos correspondientes son 3,73, 3,31 y 3,12.

El procedimiento de detección de *outliers* recomendado por Vogelsang (1999) es aplicado de la siguiente manera. Primero, el estadístico τ es calculado para la serie completa y τ es comparado al valor crítico apropiado. Si τ excede el valor crítico, entonces un *outlier* es detectado a la posición $\hat{T}_{ao} = \arg \max_{T_{ao}} |t_{\delta}(T_{ao})|$. Luego de esto, la observación relacionada con el *outlier* detectado es extraída y (2) es otra vez estimado, y se realiza una verificación por la presencia de otro *outlier*. El procedimiento continúa hasta que no se obtiene ningún rechazo.

Desafortunadamente, el procedimiento originalmente propuesto por Vogelsang (1999) tiene severas distorsiones de nivel cuando se aplica en una forma iterativa en la búsqueda de *outliers* aditivos. La explicación para este problema es el hecho de que la distribución asintótica del estadístico τ es solamente válida para la primera ronda de las iteraciones. Esto es especificado y explicado en el teorema 1 de Perron y Rodríguez (2003). Dicho teorema muestra que los valores críticos asintóticos deben ser modificados. En general, el teorema señala que los valores críticos correctos que deben ser usados en el procedimiento iterativo completo son α^i , donde α es el nivel de significancia e i es el número de la etapa en el procedimiento iterativo.

En lo que sigue, denotaremos por τ_c el procedimiento iterativo de detección de *outliers* que utiliza los correctos (y diferentes) valores críticos asintóticos en las diferentes etapas del procedimiento iterativo. En las aplicaciones empíricas

² Para ser más precisos, $t_{\delta}(T_{ao}) \Rightarrow H(\lambda) = W^*(\lambda) / (\int_0^1 W^*(r)^2 dr)^{1/2}$, donde $W^*(\lambda)$ significa un proceso de Wiener al cual se ha extraído la media. Si (2) también incluye una tendencia, $W^*(\lambda)$ significará un proceso de Wiener al cual se le ha extraído la tendencia. Adicionalmente, utilizando el teorema de mapeo continuo, obtenemos que $\tau = \sup_{\lambda \in (0,1)} |H(\lambda)| \equiv H^*$. Esta distribución es invariante con respecto a cualquier parámetro de ruido, incluida la estructura de correlación de la función de ruido.

utilizaremos valores críticos correspondientes a 5,0% de nivel de significancia. Esos valores críticos son 2,989, 3,689, 4,294 y 4,425 para la primera, segunda, tercera y cuarta etapa del procedimiento iterativo.

Un problema con el método descrito anteriormente es que su potencia es débil a menos que el tamaño del *outlier* sea largo.³ Perron y Rodríguez (2003) han propuesto una estrategia iterativa con más potencia utilizando un estadístico basado en las primeras diferencias de los datos. Considérese el proceso generador de datos dado por (1) con $d_t = \mu$, y un solo *outlier* ocurre a la fecha T_{ao} con magnitud δ . Entonces,

$$\Delta y_t = \delta [D(T_{ao})_t - D(T_{ao})_{t-1}] + v_t \quad (3)$$

Donde $D(T_{ao})_t = 1$, si $t = T_{ao}$ (0, si otro) y $D(T_{ao})_{t-1} = 1$, si $t = T_{ao} - 1$ (0, si otro). Si los datos contienen tendencia, una constante debería ser incluida. En este caso, nuestro interés es en $\tau_d = \sup_{T_{ao}} |t_{\delta}(T_{ao})|$, donde $t_{\delta}(T_{ao}) = \hat{\delta} / 2(\hat{R}_u(0) - \hat{R}_u(1))$ y $R_u(j)$ es la función de autocovarianza de u_t al retardo j .⁴

Para detectar múltiples *outliers*, seguimos una estrategia similar a la propuesta por Vogelsang (1999); es decir, eliminamos la observación calificada como *outlier* antes de proceder a la próxima etapa. El hecho importante es que, a diferencia del caso de los tests basados en los niveles de las series (como el estadístico τ de Vogelsang), la distribución asintótica del estadístico τ_d es la misma para cada una de las etapas de las iteraciones.⁵

Siguiendo la práctica común en la literatura, Perron y Rodríguez (2003) han tabulado valores críticos asumiendo errores *i. i. d.* normales para $T = 100$ y $T = 200$. Dado que en el presente documento disponemos de series más largas, hemos simulado valores críticos para $T = 250$ y $T = 350$ para los dos posibles casos de componentes determinísticos ($z_t = \{1\}$ y $z_t = \{1, t\}$) y fueron utilizadas 5.000 replicaciones. Los puntos porcentuales del estadístico τ_d son presentados en la tabla 1, que también incluye los valores críticos para $T = 100$ y $T = 200$.

En la literatura existen dos sugerencias relacionadas con la verificación de una raíz unitaria en la presencia de *outliers* aditivos. Una de ellas propone la utilización de estadísticos de raíz unitaria robustos a la presencia de fuerte correlación de

³ Evidencia basada en simulaciones referentes a este punto pueden ser halladas en Perron y Rodríguez (2003).

⁴ Donde $\hat{R}_u(j) = T^{-1} \sum_{t=1}^{T-j} \hat{v}_t \hat{v}_{t-j}$ con \hat{v}_t los residuos de la regresión por MCO obtenidos de (3). Entonces, $\hat{R}_u(j)$ es un estimador consistente de $R_u(j)$.

⁵ La desventaja de este procedimiento comparado con aquel basado en los niveles de los datos es que la distribución asintótica depende de la distribución específica de los errores v_t , aunque no de la presencia de correlación serial y heterocedasticidad. Este problema es exactamente el mismo que aquel encontrado en la detección de *outliers* en series de tiempo estacionarias.

tipo media móvil en los residuos (véase Vogelsang 1999). La segunda sugerencia propone el uso del estadístico ADF corregido por variables *dummies* asociadas a las fechas de los *outliers* aditivos identificados en una etapa preliminar (véase Franses y Haldrup 1994, y Vogelsang 1999).

Estadísticos que son robustos a la presencia de correlación serial negativa de tipo media móvil son los llamados tests M , que fueron originalmente propuestos por Stock (1999) y posteriormente analizados por Perron y Ng (1996). Recientemente, Ng y Perron (2001) han mostrado que dichos estadísticos tienen una *performance* aún mejor en términos de potencia (y nivel) cuando los datos son transformados utilizando Mínimos Cuadrados Generalizados (MCG). Estos estadísticos (M^{GLS}) son definidos de la siguiente manera:

$$MZ_{\alpha}^{GLS} = (T^{-1} \tilde{y}_T^2 - s^2)(2T^{-2} \sum_{t=1}^T \tilde{y}_{t-1}^2)^{-1} \quad (4)$$

$$MSB^{GLS} = (T^{-2} \sum_{t=1}^T \tilde{y}_{t-1}^2 / s^2)^{1/2} \quad (5)$$

$$MZ_t^{GLS} = (T^{-1} \tilde{y}_T^2 - s^2)(4s^2 T^{-2} \sum_{t=1}^T \tilde{y}_{t-1}^2)^{-1/2} \quad (6)$$

Donde $\tilde{y}_t = y_t - \hat{\psi} z_t$, $\hat{\psi}$ es el estimador que minimiza la función $S^*(\psi) = \sum_{t=0}^T (y_t^{\bar{\alpha}} - \psi' z_t^{\bar{\alpha}})^2$ con $y_t^{\bar{\alpha}} = (y_1, (1 - \bar{\alpha}L)y_t)$, $z_t^{\bar{\alpha}} = (z_1, (1 - \bar{\alpha}L)z_t)$, para $t = 1, 2, 3, \dots, T$ con $\bar{\alpha}$ denotando un parámetro de no centralidad que mide la distancia respecto de la hipótesis nula. Tal como ha sido recomendado por Elliott, Rothenberg y Stock (1996), este parámetro es igual a -7,0 ó -13,5 cuando solamente un intercepto o cuando un intercepto y una tendencia son incluidos en la regresión, respectivamente. El término s^2 es un estimador autorregresivo de (2π veces) la densidad espectral a la frecuencia cero de v_t , definida como $s^2 = s_{\eta k}^2 / (1 - \hat{b}(1))^2$, donde $s_{\eta k}^2 = T^{-1} \sum_{t=k+1}^T \hat{\eta}_{tk}^2$, $\hat{b}(1) = \sum_{j=1}^k \hat{b}_j$ y $\{\hat{\eta}_{tk}\}$ son obtenidos de la siguiente autorregresión:⁶

$$\Delta \tilde{y}_t = b_0 \tilde{y}_{t-1} + \sum_{j=1}^k b_j \Delta \tilde{y}_{t-j} + \eta_{tk} \quad (7)$$

Utilizar un test ADF corregido por la presencia de *outliers* es equivalente a utilizar el t-estadístico para verificar que $\alpha = 1$ en la siguiente regresión:

$$y_t = \mu + \alpha y_{t-1} + \sum_{j=0}^{k+1} \delta_j D(T_{ao,j})_{t-j} + \sum_{i=0}^k d_i \Delta y_{t-i} + e_t \quad (8)$$

⁶ Las ventajas de utilizar este estimador autorregresivo de la densidad espectral respecto de métodos tradicionales basados en *kernels* se discuten en Perron y Ng (1998).

Donde $D(T_{ao,j})_t = 1$ si $t = T_{ao,j}$ y 0 si otro, con $T_{ao,j}$ ($j = 1, 2, \dots, m$) siendo las fechas de los *outliers* identificados. Es necesario notar que $k + 2$ variables *dummies* deben ser incluidas en (8) con el objetivo de eliminar toda posible influencia de los *outliers* aditivos.

Evidencia empírica

Nivel y potencia del estadístico ADF corregido por outliers aditivos

Se realizaron simulaciones de nivel y de potencia del estadístico ADF con y sin corrección por la presencia de *outliers* aditivos.⁷ En general, los dos procedimientos τ_c y τ_d trabajan bien en lo que respecta a la detección de *outliers* aditivos. Sin embargo, el estadístico τ_d nos permite tener mejores propiedades en términos del nivel del estadístico ADF con un nivel exacto muy cercano al nominal. Esta evidencia es válida si el retardo de la regresión es fijo o es escogido utilizando el procedimiento secuencial basado en la significancia del estadístico t tal como ha sido sugerido por Campbell y Perron (1991).⁸ Una posible limitación de este método para seleccionar el retardo es su tendencia a escoger un retardo largo aun cuando la estructura no lo amerite. De este modo, dicho método tiende a reducir artificialmente las distorsiones de nivel. En la tabla 2 se muestran resultados utilizando un retardo fijo igual a uno con errores $AR(1)$. A manera de ejemplo, nótese que con $\rho = -0,40$, τ_c y τ_d tienen niveles iguales a 0,212 y 0,055, respectivamente, comparado al caso en el que dichos estadísticos tienen niveles iguales a 0,101 y 0,046 utilizando el método secuencial.

3. Inflación en América Latina

En los últimos años, la mayoría de los países de América Latina han experimentado diferentes programas de estabilización con el objetivo de detener los episodios de alta inflación. Intervenciones de este tipo, en la mayoría de los casos, han introducido *outliers* aditivos en la evolución de las series de inflación. En este sentido, las series de inflación de América Latina ofrecen un buen ejemplo de la presencia de este tipo de observaciones en el contexto de posible no estacionariedad.

⁷ Están disponibles los detalles completos sobre las simulaciones y tablas si se los solicita al autor.

⁸ Por ejemplo, cuando tenemos innovaciones de tipo media móviles con $\theta = -0,80$ y largos *outliers* aditivos, el nivel del estadístico ADF sin corrección por *outliers* aditivos es 0,743. Los otros dos métodos tienen niveles iguales a 0,460 y 0,392, respectivamente. Cuando $\theta = 0,0$, los tres métodos tienen niveles iguales a 0,133, 0,087 y 0,041, respectivamente.

En esta sección se consideran cuatro series de inflación de América Latina. Los países seleccionados son Argentina, Bolivia, Chile y Perú. Los períodos de alta inflación en Argentina y Perú se localizan entre 1985 y 1990, cuando se aplicaron los más importantes programas de estabilización. En el caso de la Argentina, los planes de estabilización más conocidos fueron el *Programa Austral* (junio de 1985), el programa de febrero de 1987, el *Programa Austral II* (octubre de 1987), el *Programa Primavera* (agosto de 1988), el *Programa BB* (1989), el *Programa Bonex* (enero de 1990) y el programa del ministro Domingo Cavallo (marzo de 1991). Las fechas en paréntesis corresponden a las fechas de inicio de dichos programas. En el caso del Perú, pueden mencionarse dos programas de estabilización. Ellos son el programa del ministro Salinas (setiembre de 1988) y el programa del presidente Fujimori (julio-agosto de 1990). En el caso de Bolivia, el episodio de alta inflación ocurrió a mediados de 1980. Varios pequeños programas de estabilización fueron aplicados durante el período comprendido entre 1982 y 1984, pero fue el programa aplicado en agosto de 1985 el que logró detener la inflación. Finalmente, alrededor del año 1975 se produjeron varios episodios de inflación en Chile. Diversos programas fueron aplicados entre 1975 y 1977, hasta el programa de estabilización aplicado a partir de fines de 1977 hasta 1979.

Las estimaciones se basan en datos mensuales. Para la Argentina, el período de información abarca desde enero de 1979 hasta marzo de 1999. Para Bolivia y el Perú, desde enero de 1979 hasta mayo del 2000. Para Chile, desde enero de 1970 hasta mayo del 2000. Los períodos fueron elegidos teniendo en cuenta dos criterios. El primer criterio es el interés de incluir todos o los principales programas de estabilización. Por ejemplo, la estabilización en Chile fue realizada en la década de 1970, y esa es la razón por la cual el período es diferente del resto de países. El segundo criterio es la disponibilidad de la información.

Las cuatro series se presentan en la figura 1. Un análisis visual indica una clara presencia de *outliers* aditivos (asociados con los diferentes programas de estabilización mencionados anteriormente). En diferentes aplicaciones empíricas, estas series son frecuentemente modeladas como procesos $I(0)$, y esta consideración se basa en la aplicación de tests de raíz unitaria como el estadístico ADF o el Phillips-Perron. Dado que existen largos *outliers* aditivos, nuestra hipótesis es que aquellos resultados son incorrectos debido a que dichos *outliers* aditivos introducen un componente fuertemente negativo de tipo media móvil, lo cual conduce a un fuerte rechazo de la hipótesis nula de una raíz unitaria. En consecuencia, postulamos que cuando se toma en cuenta una corrección por la presencia de *outliers* aditivos, el rechazo de la hipótesis de una raíz unitaria no es posible o, en todo caso, la evidencia en contra de ella es débil.

Evidencia de tests de raíz unitaria

En esta sección aplicamos tests comunes de raíz unitaria. Al final también consideramos la aplicación de los estadísticos M^{GLS} , los cuales son considerados robustos por la presencia de correlación serial negativa de tipo media móvil.

El retardo es elegido siguiendo el procedimiento secuencial basado en el t-estadístico. Para las tablas 3 y 4 utilizamos $kmax = \text{int}[12 * (T/100)^{1/2}]$. La mayoría de los resultados van en la misma dirección; es decir, indican un fuerte rechazo de la hipótesis nula de raíz unitaria. Dichos resultados son claros incluso en presencia de largos retardos utilizados en las regresiones. Un comentario similar se obtiene cuando se aplican estadísticos de raíz unitaria que utilizan información transformada por MCG (véase la tabla 4).

Detección de outliers aditivos

La siguiente etapa es la detección de *outliers* aditivos para así verificar nuestra inspección visual de la figura 1. Dado que utilizamos datos mensuales, preferimos utilizar un valor crítico al 1,0% de nivel de significancia, con el objetivo de evitar detectar un número excesivo de *outliers* aditivos. Esto se hace para el procedimiento basado en las primeras diferencias de los datos (τ_d). Cuando se utiliza el procedimiento basado en diferentes valores críticos para diferentes etapas (τ_c), se utilizan también valores críticos al 5,0%.

Los resultados (véase la cuarta columna en las tablas 5 y 6) muestran que existe un número significativo de *outliers* aditivos en todas las series de inflación analizadas. Como ya se mencionó previamente, el estadístico τ_d es más potente y los resultados revelan este hecho. Ambos procedimientos son capaces de detectar los principales *outliers* aditivos como aquellas observaciones asociadas directamente a la aplicación de programas de estabilización (y observaciones adyacentes). Esto se refleja claramente en el caso peruano. Los tres *outliers* detectados corresponden exactamente al programa Salinas (setiembre de 1988) y al programa del presidente Fujimori (julio-agosto de 1990).

Evidencia del estadístico ADF corregido por outliers aditivos

La etapa final es la aplicación del estadístico ADF utilizando variables *dummies* que toman en cuenta la presencia de *outliers* aditivos detectados en la cuarta columna de las tablas 5 y 6. En este caso, utilizamos un $kmax = \text{int}[12 * (T/100)^{1/4}]$. Los resultados del procedimiento τ_d se presentan en la tabla 5. De acuerdo con estos resultados, las series de inflación para Argentina y Perú pueden ser consideradas

como procesos $I(1)$. En los casos de Bolivia y Chile, la inflación puede ser considerada como un proceso $I(0)$. Es necesario notar que se seleccionó un corto retardo para el caso de Bolivia y podría existir alguna duda sobre el rechazo de la hipótesis nula. Sin embargo, imponiendo un $k_{min} = 6$ o un $k_{min} = 12$, los t -estadísticos son -3,69 y -2,62, los cuales son significativos al 1,0% y al 10,0% de significancia, respectivamente.

Los resultados obtenidos del procedimiento τ_c se presentan en la tabla 6. Estos son similares a aquellos obtenidos en la tabla 5, a excepción del caso de Bolivia, que presenta evidencia de una raíz explosiva. Aunque este resultado es diferente del obtenido en la tabla 5, se obtiene el rechazo de la hipótesis nula de una raíz unitaria, lo cual es importante aquí.

Es necesario hacer un comentario relacionado con el número de *outliers* detectados utilizando el método τ_c . Desafortunadamente, solo podemos utilizar cuatro valores críticos (utilizando el procedimiento τ_c), los cuales corresponden a la primera, segunda, tercera y cuarta etapas del procedimiento iterativo para detectar *outliers* aditivos. La razón por la cual podemos utilizar únicamente este limitado número de valores críticos es el hecho de que cada vez necesitamos valores críticos muy lejanos en las colas de la distribución (véase Perron y Rodríguez 2003). Una posible alternativa ad hoc es el uso del valor crítico correspondiente a la cuarta etapa para todas las etapas subsecuentes. Esto se hizo hasta un máximo de 36 iteraciones. El número de *outliers* detectados es 36; similares resultados se obtuvieron respecto de aquellos mostrados en la tabla 6.

Lucas (1995a, 1995b) propone el uso de estimados robustos a la presencia de *outliers*. El procedimiento no es secuencial, lo cual puede ser una ventaja en ciertos casos en los que la identificación de las fechas de los *outliers* no es de interés. Sin embargo, en la mayoría de los casos, el investigador está interesado en conocer las fechas de dichas observaciones. Este es nuestro caso, entre otras razones porque la inspección visual es verificada, y desde una perspectiva macroeconómica, la identificación de *outliers* nos permite identificar los eventos que los originaron.

Cuando las fechas de los *outliers* son necesarias, el procedimiento recomendado por Lucas (1995a, 1995b) necesita la especificación de un valor crítico generalmente localizado entre 3 y 4. En este caso, nuestros resultados verifican el hecho de que la recomendación de Lucas (1995a, 1995b) adolece del mismo problema que Vogelsang (1999), dado que un excesivo número de observaciones se clasifican como *outliers*. Sin embargo, el estimado relacionado con el coeficiente autorregresivo en la regresión ADF puede ser considerado robusto a la presencia de *outliers*, lo que constituye el argumento de Lucas (1995a). Él utiliza los así llamados estimadores *high breakdown point* (*HBP*), propuestos originalmente por Rousseeuw y Leroy (1987). Existen varios estimadores *HBP* propuestos en la

literatura. Aquí utilizamos el estimador S (Rousseeuw y Yohai 1984), el estimador MM (Yohai 1987).⁹ Por otro lado, Lucas (1995b) propone un estadístico para evaluar la existencia de una raíz unitaria sobre la base del estimador M (Hampel *et al.* 1986; Huber 1981). Este tipo de estimador es conocido por tener cierto grado de insensibilidad a la presencia de *outliers*. En este procedimiento, se necesita la estimación de una función particular $\psi(\cdot)$. Algunos detalles y supuestos que tienen que ser satisfechos por esta función se mencionan en Lucas (1995b). Aquí utilizamos dos opciones particulares para la función $\psi(\cdot)$, las cuales son la función *bi-square* $\psi(\cdot)$ y la función Huber $\psi(\cdot)$ (Huber 1981); véase Lucas (1995b) para detalles adicionales.

En la aplicación de los procedimientos, se utilizó un retardo similar a los casos de τ_d y τ_c . Para el estimador S , se utilizaron 10.000 replicaciones en la especificación del muestreo aleatorio. Mientras que varios *outliers* son hallados para cada una de las series de inflación, en general, los resultados que utilizan los estimadores HBP y M confirman nuestros resultados previos. En el caso de la Argentina, la hipótesis nula no es rechazada utilizando el estimador MM ($t_\alpha = -0,31$). El estimador M soporta la misma conclusión ($t_\alpha = -0,83$ y $t_\alpha = -1,27$). Solamente el estimador S confirma los resultados obtenidos por MCO; es decir, un rechazo de la hipótesis nula ($t_\alpha = -7,49$). Sin embargo, es conocido en la literatura que este tipo de estimador es sesgado cuando existen *outliers* en las variables explicativas, lo que se llama con frecuencia efecto *leverage*. En el caso de Bolivia, todos los procedimientos sugieren un fuerte rechazo de la hipótesis nula de una raíz unitaria ($t_\alpha = -15,41$ utilizando el estimador MM ; $t_\alpha = -65,37$ utilizando el estimador S ; y $t_\alpha = -17,96$ usando el estimador M con una función *bi-squared* $\psi(\cdot)$). La sola excepción es el estimador M cuando una función Huber $\psi(\cdot)$ es utilizada ($t_\alpha = -1,11$). En el caso de la inflación en Chile, todos los procedimientos rechazan la hipótesis de raíz unitaria ($t_\alpha = -4,09$ usando el estimador MM ; $t_\alpha = -7,83$ utilizando el estimador S ; y $t_\alpha = -4,91$ usando el estimador M con una función *bi-squared* $\psi(\cdot)$). El estimador M utilizando la función Huber $\psi(\cdot)$ proporciona un resultado de significancia muy cercano a 10,0%. Finalmente, la inflación en el Perú es estacionaria de acuerdo con los resultados del estimador MM y S ($t_\alpha = -3,25$ y $t_\alpha = -7,28$, respectivamente). Sin embargo, el estimador M (con ambos tipos de funciones $\psi(\cdot)$) indica que la inflación es no estacionaria ($t_\alpha = -1,32$ y $-0,73$, respectivamente).

⁹ No se presentan los resultados obtenidos de utilizar el estimador de la mediana mínima de cuadrados (*Least Median of Squares*, *LMS*), sugerido por Rousseeuw (1984), dado que los errores estándar no están disponibles. La razón es que dicho estimador tiene propiedades asintóticas pobres. Véase Lucas (1995a, 1995b).

4. Conclusiones

Utilizar cuatro series de inflación de cuatro países de América Latina ha permitido ilustrar los peligros empíricos del uso de estadísticos conocidos para detectar una raíz unitaria. De hecho, la aplicación de los estadísticos M y ADF utilizando datos transformados por MCG han mostrado un fuerte rechazo de la hipótesis de raíz unitaria. Dada la fuerte magnitud de los *outliers* aditivos presentes en las cuatro series de inflación, se recomienda y utiliza el estadístico ADF corregido por la presencia de *outliers*. Los resultados sugieren que las series de inflación en la Argentina y el Perú pueden ser consideradas como procesos $I(1)$. Para los casos de Bolivia y Chile, dichas series pueden considerarse estacionarias.

Nuestros resultados han sido comparados con aquellos obtenidos de los procedimientos propuestos por Lucas (1995a, 1995b). En general, los resultados son muy similares. Mientras que los procedimientos sugeridos por Lucas (1995a, 1995b) no son secuenciales, lo cual es una ventaja, adolecen del mismo defecto que el procedimiento de Vogelsang (1999), el cual consiste en la identificación de un número excesivo de *outliers*. Cuando el investigador está solamente interesado en la estimación robusta de un coeficiente en particular, una buena opción es la aplicación de los procedimientos sugeridos por Lucas (1995a, 1995b). Sin embargo, cuando la identificación de las fechas de ocurrencia de los *outliers* es importante, es aconsejable el procedimiento propuesto por Perron y Rodríguez (2003).

5. Apéndice: fuentes de información

Las series de inflación de Perú provienen del Instituto Nacional de Estadística e Informática y del Banco Central de Reserva. Guillermo Díaz nos proporcionó gentilmente dicha información. Para el caso de Chile, la fuente es el Instituto Nacional de Estadística. Fue Rafael Herrada quien nos proporcionó dicha información, mientras se encontraba siguiendo sus estudios de maestría en Chile. Para Bolivia, los datos hasta diciembre de 1998 fueron gentilmente otorgados por Wilfredo Sillerico Gálvez, del Banco Central de Bolivia. El resto de la información se completó a partir de la página web del Banco Central de Bolivia. La serie de inflación de la Argentina para el período enero de 1979-diciembre de 1996 es una submuestra de la muestra total utilizada por Baillie (1996). El profesor Baillie nos proporcionó dicha información. El resto de la información ha sido completada a través de la página web del Instituto Nacional de Estadística de Argentina.

Tabla 1. Valores críticos finitos para τ_d

Significancia		$z_t = \{1\}$					$z_t = \{1,t\}$		
	$T = 100$	$T = 200$	$T = 250$	$T = 350$	$T = 100$	$T = 200$	$T = 250$	$T = 350$	
1,0%	4,13	4,15	4,22	4,27	4,11	4,14	4,21	4,27	
2,5%	3,86	3,95	3,95	4,04	3,83	3,94	3,94	4,04	
5,0%	3,65	3,78	3,79	3,86	3,64	3,77	3,78	3,85	
10,0%	3,44	3,56	3,61	3,66	3,42	3,55	3,60	3,66	

Tabla 2. Nivel del estadístico ADF, errores AR(1) - (retardo igual a uno)

		$\delta_1 = 0, \delta_2 = 0,$			$\delta_1 = 5, \delta_2 = 3,$			$\delta_1 = 10, \delta_2 = 5,$		
		$\delta_3 = 0, \delta_4 = 0,$			$\delta_3 = 2, \delta_4 = 2,$			$\delta_3 = 5, \delta_4 = 5,$		
		nc	τ_c	τ_d	Nc	τ_c	τ_d	nc	τ_c	τ_d
$\rho = -0,80$	sin	0,052	0,031	0,050	0,169	0,043	0,100	0,502	0,006	0,012
	con	0,000	0,024	0,002	0,000	0,081	0,033	0,000	0,252	0,175
	total	0,052	0,055	0,052	0,169	0,124	0,133	0,502	0,258	0,187
$\rho = -0,40$	sin	0,052	0,027	0,046	0,130	0,038	0,008	0,380	0,009	0,000
	con	0,000	0,028	0,005	0,000	0,062	0,070	0,000	0,203	0,055
	total	0,052	0,055	0,051	0,130	0,100	0,078	0,380	0,212	0,055
$\rho = 0,00$	sin	0,057	0,031	0,055	0,070	0,029	0,000	0,198	0,017	0,000
	con	0,000	0,023	0,002	0,000	0,039	0,057	0,000	0,103	0,051
	total	0,057	0,054	0,057	0,070	0,068	0,057	0,198	0,120	0,051
$\rho = 0,40$	sin	0,061	0,038	0,059	0,034	0,017	0,000	0,054	0,011	0,000
	con	0,000	0,017	0,001	0,000	0,015	0,045	0,000	0,032	0,044
	total	0,061	0,055	0,060	0,034	0,032	0,045	0,054	0,043	0,044
$\rho = 0,80$	sin	0,060	0,054	0,058	0,027	0,021	0,000	0,036	0,028	0,000
	con	0,000	0,007	0,002	0,000	0,006	0,039	0,000	0,008	0,054
	total	0,060	0,061	0,060	0,027	0,027	0,039	0,036	0,036	0,054

El proceso generador de datos es $y_t = \sum_{j=1}^4 \delta_j D(T_{ao,j})_t + u_t$ con $u_t = u_{t-1} + v_t$ y $v_t = \rho v_t + \epsilon_t$. La columna denotada por «nc» es el nivel del estadístico ADF sin corrección por la presencia de *outliers*. La columna denotada por τ_c y τ_d es el nivel del estadístico ADF corregido por la presencia de *outliers* utilizando los procedimientos τ_c y τ_d respectivamente. La línea denotada por «sin» muestra el

nivel del estadístico *ADF* cuando no se ha realizado ninguna corrección por la presencia de *outliers*. La línea denotada por «con» muestra el nivel del estadístico *ADF* cuando se ha hecho una corrección por la presencia de *outliers*. Finalmente, la línea denotada por «total» es simplemente la suma de las dos líneas anteriormente descritas.

Tabla 3. Estadísticos *ADF* y *Phillips-Perron*
(La regresión incluye solo una constante)

País	Muestra	kmax	t_{α}	α	k	Z_t	k
Argentina	1979:01-1999:03	19	-2,809 ^d	0,796	15	-7,205 ^a	1
Bolivia	1979:01-2000:05	19	-2,932 ^c	0,822	10	-8,83 ^a	10
Chile	1970:01-2000:05	23	-2,050	0,881	21	-16,53 ^a	21
Perú	1979:01-2000:05	19	-5,300 ^a	0,512	3	-12,32 ^a	3

a, b, c, d denotan niveles de significancia al 1,0%, 2,5%, 5,0% y 10,0%, respectivamente.

Tabla 4. Estadísticos M^{GLS} , P_T^{GLS} y ADF^{GLS}
(Componentes determinísticos $z_t = \{1\}$ $c = -7.0$)

País	Período	kmax	$MZ\alpha$	MSB	MZ_t	P_T	ADF	α	K
Argentina	1979:01-1999:03	19	-30,96 ^a	0,126 ^a	-3,922 ^a	0,834 ^a	-2,748 ^a	0,804	15
Bolivia	1979:01-2000:05	19	-18,87 ^a	0,163 ^a	-3,071 ^a	-1,367 ^a	-2,734 ^a	0,846	10
Chile	1970:01-2000:05	23	-9,85 ^b	0,224 ^c	-2,208 ^b	-2,519 ^b	-2,052 ^c	0,880	21
Perú	1979:01-2000:05	19	-40,17 ^a	0,111 ^a	-4,480 ^a	0,609 ^a	-5,283 ^a	0,525	3

a, b, c, d denotan niveles de significancia al 1,0%, 2,5%, 5,0% y 10,0%, respectivamente.

Los valores críticos para los estadísticos M^{GLS} se obtienen de Ng y Perron (2001).

Los valores críticos para el estadístico ADF^{GLS} son equivalentes al caso en el que existen no componentes determinísticos, como ha sido demostrado por Elliott, Rothenberg y Stock (1996).

Tabla 5. Estadístico *ADF* corregido por outliers aditivos utilizando τ_d
(Componentes determinísticos $z_t = \{1\}$)*

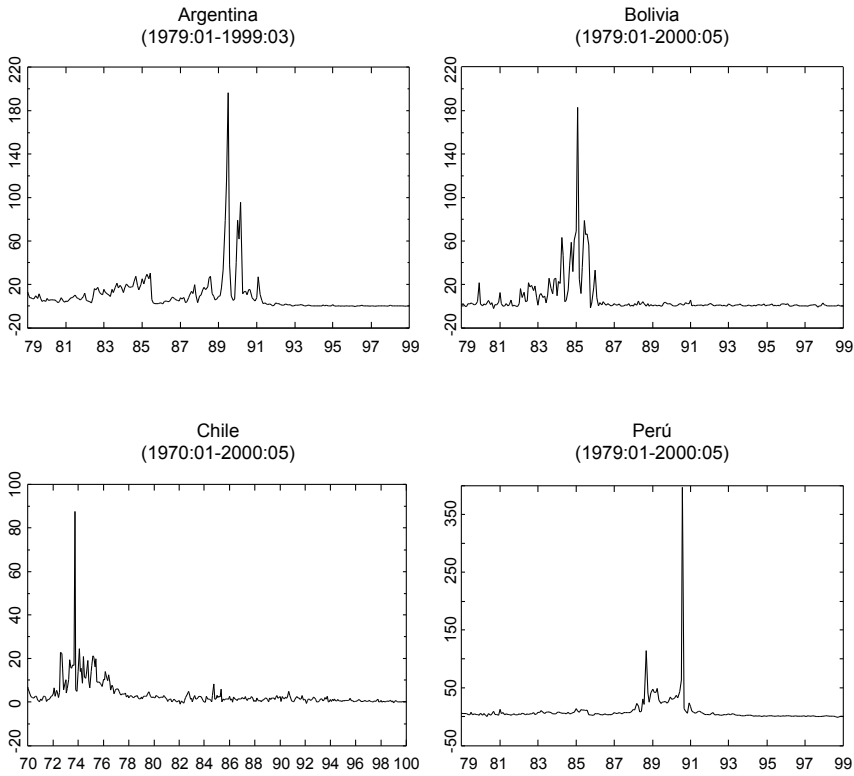
País	Muestra	kmax	Outliers	t_{α}	α	k
Argentina	1979: 01-1999:03	15	12	2,424	1,104	14
Bolivia	1979: 01-2000:05	15	30	-4,570 ^a	0,399	5
Chile	1970: 01-2000:05	17	19	-4,683 ^a	0,830	13
Perú	1979: 01-2000:05	15	10	-0,652	0,968	10

* Los *outliers* aditivos fueron detectados utilizando valor crítico al 1,0%.

Tabla 6. Estadístico ADF corregido por outliers aditivos utilizando τ_c
 (Componentes determinísticos $z_t = \{1\}$)

País	Muestra	kmax	Outliers	t_α	α	k
Argentina	1979: 01-1999:03	15	4	0,128	1,006	14
Bolivia	1979: 01-2000:05	15	4	4,049	1,289	12
Chile	1970: 01-2000:05	17	4	-7,139 ^a	0,852	11
Perú	1979: 01-2000:05	15	4	0,151	1,009	13

Figura 1. Inflación mensual en América Latina



Referencias bibliográficas

- BAILLIE, R. T. (1996) «Long Memory Process and Fractional Integration in Econometrics». *Journal of Econometrics*, vol. 73, n.º 1, pp. 5-59.
- CAMPBELL, J. Y. y P. PERRON (1991) «Pitfalls and Opportunities: What Macroeconomists Should Know About Unit Roots». En O. J. Blanchard y S. Fisher (editores). *NBER Macroeconomic Annual*, vol. 6, pp. 141-201.
- CHANG, I., G. C. TIAO y C. CHEN (1988) «Estimation of Time Series Parameters in the Presence of Outliers». *Technometrics*, n.º 30, pp. 193-204.
- CHEN, C. y L. LIU (1993) «Joint Estimation of Model Parameters and Outlier Effects in Time Series». *Journal of the American Statistical Association*, n.º 74, pp. 427-431.
- DICKEY, D. A. y W. A. FULLER (1979) «Distribution of the Estimators for Autoregressive Time Series with a Unit Root». *Journal of the American Statistical Association*, n.º 74, pp. 427-431.
- ELLIOTT, G., T. ROTHENBERG y J. H. STOCK (1996) «Efficient Tests for an Autoregressive Unit Root». *Econometrica*, n.º 64, pp. 813-836.
- FRANSES, P. H. y N. HALDRUP (1994) «The Effects of Additive Outliers on Tests for Unit Roots and Cointegration». *Journal of Business & Economic Statistics*, n.º 12, pp. 471-478.
- GÓMEZ, V. y A. MARAVALL (1992a) «Estimation, Prediction and Interpolation for Nonstationary Series with the Kalman Filter». European University Institute, Working Paper ECO 92/80.
- (1992b) «Time Series Regression with ARIMA Noise and Missing Observations. Program TRAM». European University Institute, Working Paper ECO 92/81.
- HAMPEL, F. R., E. M. RONCHETTI, P. J. ROUSSEEUX y W. A. STAHEL (1986) *Robust Statistics: The Approach Based on Influence Functions*. Nueva York: Wiley.
- HOEK, H., A. LUCAS y H. K. VAN DIJK (1995) «Classical and Bayesian Aspects of Robust Unit Root Inference». *Journal of Econometrics*, n.º 69, pp. 27-59.
- HUBER, P. J. (1981) *Robust Statistics*. Nueva York: Wiley.
- LUCAS, A. (1995a) «An Outlier Robust Unit Root Test with an Application to the Extended Nelson-Plosser Data». *Journal of Econometrics*, n.º 66, pp. 153-173.
- (1995b) «Unit Root Tests Based on M Estimators». *Econometric Theory*, n.º 11, pp. 331-346.
- NG, S. y P. PERRON (2001) «Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power». Por publicarse en *Econometrica*.
- PERRON, P. y S. NG (1998) «An Autoregressive Spectral Density Estimator at Frequency Zero for Nonstationarity Tests». *Econometric Theory*, n.º 14, pp. 560-603.
- (1996) «Useful Modifications to Unit Root Tests with Dependent Errors and their Local Asymptotic Properties». *Review of Economic Studies*, n.º 63, pp. 435-463.

- PERRON, P. y G. RODRÍGUEZ (2003) «Searching for Additive Outliers in Nonstationary Time Series». Por publicarse en *Journal of Time Series Analysis*.
- PHILLIPS, P. C. B. (1987) «Time Series Regression with Unit Roots». *Econometrica*, n.º 55, pp. 277-302.
- PHILLIPS, P. C. B. y P. PERRON (1988) «Testing for a Unit Root in Time Series Regression». *Biometrika*, n.º 75, pp. 335-346.
- ROUSSEEUW, P. J. (1984) «Least Median of Squares Regression». *Journal of the American Statistical Association*, n.º 79, pp. 871-880.
- ROUSSEEUW, P. J. y A. M. LEROY (1987) *Robust Regression and Outliers Detection*. Nueva York: Wiley.
- ROUSSEEUW, P. J. y V. J. YOHAI (1984) «Robust Regression by Means of S-Estimators». En Franke, J., W. Hardle y R. D. Martin (editores). *Robust and Nonlinear Time Series*. Nueva York: Springer, pp. 256-272.
- SAID, S. E. y D. A. DICKEY (1984) «Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order». *Biometrika*, n.º 71, pp. 599-608.
- SHIN, D. W., S. SHARKAR y J. H. LEE (1996) «Unit Root Tests for Time Series with Outliers». *Statistics and Probability Letters*, n.º 30, pp. 189-197.
- STOCK, J. H. (1999) «A Class of Tests for Integration and Cointegration». En Engle, R. F. y H. White (editores). *Cointegration, Causality and Forecasting. A Festschrift in Honour of Clive W. J. Granger*. Oxford University Press, pp. 137-167.
- TSAY, R. S. (1986) «Time Series Model Specification in the Presence of Outliers». *Journal of the American Statistical Association*, n.º 81, pp. 132-141.
- VOGELSANG, T. J. (1999) «Two Simple Procedures for Testing for a Unit Root when there are Additive Outliers». *Journal of Time Series Analysis*, n.º 20, pp. 237-252.
- YOHAI, V. J. (1987) «High Breakdown-Point and High Efficiency Robust Estimates for Regression». *Annals of Statistics*, n.º 15, pp. 642-656.