

C.E.M.A.

Virrey del Pino 3210
Belgrano R.
1426 Buenos Aires

Te. 783-3291/9311.

ANALISIS CAUSAL ENTRE DINERO Y PRECIOS:

UN ENFOQUE ROBUSTO

Roque B. Fernández

y
Víctor J. Yohai
Agosto 1980

N° 16

ANALISIS CAUSAL ENTRE DINERO Y PRECIOS:
UN ENFOQUE ROBUSTO,

por

Roque B. Fernández.
C.E.M.A.

y

Víctor J. Yohai*
C.E.M.A.

SINTESIS

En este trabajo se estudia la relación dinero-precios mediante el análisis de series de tiempo. A través de un enfoque robusto se estudian los problemas de identificación y estimación de procesos ARIMA, como así también los tests de causalidad. Los métodos robustos se utilizan para analizar la conducta de las series de Argentina en el período 1970-1978 observándose que no es posible rechazar la hipótesis de dinero exógeno (dinero causa precios). Este resultado es exactamente el opuesto cuando se analizan las series por métodos clásicos.

* Los programas de cómputo utilizados fueron desarrollados con la colaboración de la computadora científica Ester Lagomarsino del C.E.M.A. y las Licenciadas Silvia Fasano y Norma Chhab de la Comisión de Investigaciones de la Provincia de Buenos Aires a las cuales se agradece la colaboración prestada. También se agradece la colaboración de la Señorita Carola Pessino.

I. Introducción.

Las series de tiempo que frecuentemente se usan en las ciencias sociales, y en especial en el análisis económico, muy a menudo presentan observaciones con características particulares que dificultan el proceso de identificar, estimar e interpretar un modelo. Ejemplo típico de esta situación la constituye las series que miden la evolución de los precios en Argentina que, considerando un período suficientemente prolongado, muestran lapsos donde ocurrieron devaluaciones abruptas, controles de precios, "mercados negros" para muchos bienes -y su permanente omisión de las estadísticas oficiales-, liberalización repentina de precios, "pautas" de precios, etc. La presencia de observaciones anómalas, tanto la serie de inflación como en la serie de la tasa de crecimiento monetaria, se pueden observar claramente en el Gráfico N° 1, especialmente en los años 1975 y 1976. Si bien es cierto que sería altamente deseable modelar la economía de forma tal que todos estos acontecimientos pudieran ser explicados (variables endógenas), o pudieran ser incorporados como variables exógenas, también es cierto que muchas veces esto no es posible hacer si se desean mantener representaciones relativamente simple de la estructura de la economía.

Los procedimientos, clásicos, por ejemplo, los propuestos en Box and Jenkins (1970), utilizados en las diferentes etapas del análisis estadístico de series de tiempo (identificación del modelo, estimación de los parámetros, test de bondad de ajuste, tests de causalidad entre dos series) son eficientes bajo las siguientes hipótesis:

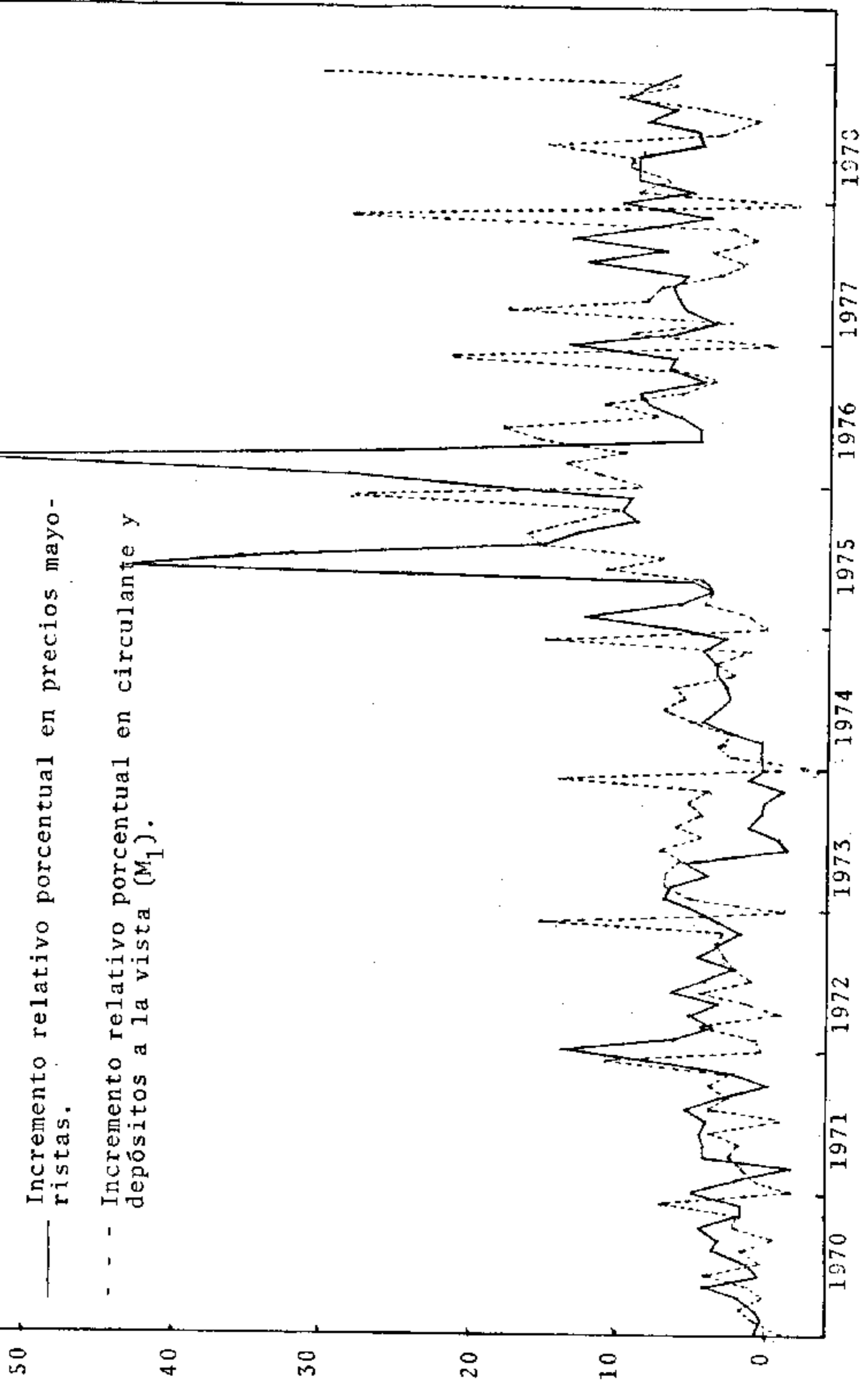
I. Introducción.

Las series de tiempo que frecuentemente se usan en las ciencias sociales, y en especial en el análisis económico, muy a menudo presentan observaciones con características particulares que dificultan el proceso de identificar, estimar e interpretar un modelo. Ejemplo típico de esta situación la constituye las series que miden la evolución de los precios en Argentina que, considerando un período suficientemente prolongado, muestran lapsos donde ocurrieron devaluaciones abruptas, controles de precios, "mercados negros" para muchos bienes -y su permanente omisión de las estadísticas oficiales-, liberalización repentina de precios, "pautas" de precios, etc. La presencia de observaciones anómalas, tanto la serie de inflación como en la serie de la tasa de crecimiento monetaria, se pueden observar claramente en el Gráfico N° 1, especialmente en los años 1975 y 1976. Si bien es cierto que sería altamente deseable modelar la economía de forma tal que todos estos acontecimientos pudieran ser explicados (variables endógenas), o pudieran ser incorporados como variables exógenas, también es cierto que muchas veces esto no es posible hacer si se desean mantener representaciones relativamente simple de la estructura de la economía.

Los procedimientos, clásicos, por ejemplo, los propuestos en Box and Jenkins (1970), utilizados en las diferentes etapas del análisis estadístico de series de tiempo (identificación del modelo, estimación de los parámetros, test de bondad de ajuste, tests de causalidad entre dos series) son eficientes bajo las siguientes hipótesis:

Gráfico N°1

— Incremento relativo porcentual en precios mayo-
ristas.
- - - Incremento relativo porcentual en circulante y
depósitos a la vista (M_1).



- (i) Todas las observaciones satisfacen el modelo.
- (ii) La distribución de los errores es normal.

Es bien conocido, ver por ejemplo Denby y Martin (1979), Yohai (1980) que estos procedimientos son muy sensitivos a pequeñas desviaciones de estas hipótesis.

Unas pocas observaciones atípicas pueden tener una enorme influencia en los resultados del análisis estadístico. Además estos procedimientos pueden ser muy ineficientes para pequeñas desviaciones de la normalidad.

Tradicionalmente la forma más simple de tratar las "perturbaciones" en el análisis econométrico de series de tiempo ha sido, por un lado, seleccionar la longitud de la muestra, tratando en lo posible evitar períodos con fuerte perturbaciones, y por otro lado, adicionar variables ficticias (dummies) para contrarrestar el efecto de las perturbaciones. Obviamente estos procedimientos son de naturaleza ad-hoc y raramente se dan otras explicaciones que no sean el probable efecto de "cambios institucionales" o de "cambios de estructura".

En este trabajo se propone un enfoque mas sistemático para el análisis estadístico de series de tiempo con observaciones atípicas. Este enfoque consiste en la utilización de procedimientos estadísticos robustos. Los procedimientos robustos son altamente eficientes bajo las hipótesis (i) y (ii) pero también son estables en presencia de unas pocas observaciones atípicas. Además los procedimientos robustos continúan siendo eficientes cuando la distribución de los errores sufre pequeñas desviaciones de la normalidad.

En este trabajo se utilizan procedimientos robustos en las siguientes etapas del análisis estadístico: identificación, estimación de parámetros de un modelo ARIMA, test de bondad de ajuste, y tests de causalidad entre dos series. Los aspectos metodológicos del problema de robustez se discuten en el Apéndice de este trabajo, mientras que, en la segunda sección, se lleva a cabo una aplicación con series de dinero y precios para Argentina en el período 1970-1978. En particular se hace uso de métodos robustos para verificar algunas hipótesis de causalidad en modelo macroeconómicos simples.

II. Análisis de las Series Dinero y Precios en Argentina; 1970-1978.

La relación entre dinero y precios es un tópico que ha sido objeto de numerosos estudios tanto en Argentina como en muchos otros países, y aunque de acuerdo a los análisis que se concentran en el largo plazo pareciera existir entre los economistas un considerable consenso sobre la naturaleza de esta relación, tal consenso parece no estar presente en las discusiones sobre corto plazo. Con tal motivo uno de los autores, en un trabajo previo (Fernández (1979)), estudió el desempeño de modelos alternativos para explicar y predecir los movimientos de corto plazo en algunas principales variables económicas. En tal oportunidad el análisis empírico se llevó a cabo utilizando métodos no robustos, aquí se verá como la aplicación de estos métodos puede provocar un cambio radical en las conclusiones.

El modelo que se analiza en este trabajo supone que la autoridad monetaria mantiene control sobre la cantidad nominal de dinero

que existe en la economía. Para ello usualmente se requiere que el sector externo funcione con un tipo de cambio libre, de manera tal que el movimiento de reservas internacionales, y de aquí la base monetaria, estén bajo control del Banco Central. En la práctica el tipo de cambio libre no es estrictamente necesario, las autoridades monetarias pueden mantener control sobre las reservas internacionales fijando el tipo de cambio por cierto período de tiempo, y recurriendo tanto a restricciones al comercio (cuotas y tarifas por ejemplo) como el movimiento de capitales internacionales (prohibición a la entrada de capitales, depósitos previos, impuestos, etc.). La oferta monetaria se deriva a partir de la base monetaria (H) que está compuesta de crédito interno (C) y reservas internacionales (R).

$$1) H \equiv C+R.$$

La base monetaria por el multiplicador (V) da la oferta monetaria:

$$2) M^S = VH.$$

Expresando (2) en tasas de cambio, y usando (1), se tiene que:

$$3) m^S = v+\alpha c+(1-\alpha)r,$$

donde la letra minúscula representa la variable original en tasa de cambio en lugar de niveles, y $\alpha = C/H$ es la participación del crédito interno en la base monetaria.

La segunda relación fundamental de este modelo es la demanda por dinero, M^d ,

$$4) M^d = AYe^{-\gamma_0 i},$$

donde Y representa el ingreso nominal, i la tasa nominal de interés, y A, γ_0 , parámetros a determinar. Diferenciando logarítmicamente (4) se tiene que:

$$5) m^d = y - \gamma_0 [1 - B] i.$$

En (5), B representa el operador de rezagos: $B_x^n(t) = x(t-n)$. Para utilizar (5) se presenta el problema de que la tasa nominal de interés ha sido una variable sujeta permanentemente a controles en las últimas décadas, estos controles se levantaron recientemente a partir de la reforma financiera de 1977. Aquí haremos uso de la ecuación de Fisher para sustituir la tasa nominal de interés:

$$6) i = \rho + p^*,$$

donde p^* es la tasa de inflación esperada, y ρ la tasa real de interés ex-ante, que se supone constante. p^* se supone que se forma en base a la historia de la inflación, es decir: $p^* = f(Bp, B^2p, B^3p, \dots)$ y que tal historia se puede captar con una función $\gamma_1(B)$ que representa a un polinomio en el operador de rezagos.

$$7) p^* = \gamma_1(B)p.$$

Sustituyendo (7) y (6) en (5) tenemos:

$$8) m^d = y - \gamma(B)p,$$

donde $\gamma(B) = \gamma_1(B)\gamma_0[1-B]$.

La última relación que necesitamos postula equilibrio en el mercado monetario:

$$9) m^d = m^s = m.$$

Dado que en la parte empírica se trabaja con datos mensuales se ignorarán los cambios en el ingreso real que pueden tener lugar mes a mes, este supuesto permite escribir:

$$10) p \approx y.$$

Luego, todo este modelo puede reducirse a:

$$3) m = v + \alpha c + (1 - \alpha)r,$$

$$11) p[1 - \gamma(B)] = m, \quad \text{ó} \quad p = [1/1 - \gamma(B)]m.$$

Entonces este modelo trabaja de la siguiente forma: a) la tasa de expansión monetaria la determina el Banco Central a través de sus operaciones con crédito interno y con reservas internacionales, b) la expansión monetaria determina la tasa de inflación a través de la estructura de rezagos.

A continuación se analizará las series mensuales disponibles para verificar si el modelo representado por las ecuaciones (3) y (11) se adapta a la experiencia argentina en el período 1970-1978. La elección de este período ha tenido en cuenta dos aspectos principales: un aspecto es contar con un período lo suficientemente corto para minimizar la cantidad de cambios institucionales que afectan la estructura de la economía; y el otro aspecto es que, utilizando datos mensuales, aún en un período corto como 1970 y 1978, se puede disponer de más de 100 observaciones para análisis estadístico. El método que utilizaremos para evaluar el modelo representado

por las ecuaciones (3) y (11) hará uso de los test de causalidad desarrollados por Pierce y Haugh. Brevemente la metodología consiste en lo siguiente: en un primer lugar se filtran las series utilizando modelos ARIMA; y en segundo lugar se encuentran las correlaciones cruzadas para someter a test la dirección de causalidad implicada por el modelo.

III. Análisis Empírico.

Para analizar las series de dinero, circulante y depósitos a la vista según datos de FMI, y precio mayoristas según datos de INDEC se utilizará simultáneamente los métodos convencionales y los métodos robustos descritos en el Apéndice. El análisis comprende tres etapas: primero, identificación de modelos ARIMA; segundo, estimación de modelos ARIMA; y tercero, análisis de causalidad.

En la Tabla 1 se muestra el correlograma para identificar el modelo para precios. Esta serie muestra estacionaridad con una primera diferencia regular del incremento relativo en el índice de precios mayoristas. El correlograma no robusto muestra autocorrelaciones significativas a rezagos, 3, 4 y 9; mientras que el correlograma robusto muestra autocorrelaciones significativas con rezagos 1 y 2. A partir de esta información se estiman algunos modelos alternativos y se estudian residuos para lograr una representación adecuada de la conducta de la serie. Los modelos seleccionados se muestran en la Tabla 3 donde se puede observar que con estimación no robusta se logra un modelo ARIMA (9,3), mientras que con estimación robusta se logra un modelo ARIMA (0,2). El estadígrafo Q tan-

Tabla N° 1

Correlograma para Precios Mayoristas.

		R E Z A G O S											
		1	2	3	4	5	6	7	8	9	10	11	12
<u>Estimación no Robusta.</u>	Primera Diferencia	-0.05	-0.28	-0.19	-0.03	0.04	0.10	-0.00	0.08	0.36	-0.13	-0.23	-0.02
	mento Relativo.												
		13	14	15	16	17	18	19	20	21	22	23	24
		R E Z A G O S											
		0.08	0.08	-0.09	-0.05	0.01	-0.01	0.14	-0.09	-0.03	0.04	0.00	-0.08
<u>Estimación Robusta.</u>	Primera Diferencia	-0.28	-0.18	-0.05	-0.11	0.16	-0.11	0.03	-0.02	0.14	-0.07	-0.17	-0.01
	Regular de Incremento Relativo.												
		13	14	15	16	17	18	19	20	21	22	23	24
		R E Z A G O S											
		0.02	0.05	0.11	-0.02	-0.17	0.09	-0.14	0.14	0.05	-0.21	0.20	-0.06

Tabla N° 2

Correlograma para Dinero.

		R E Z A G O S											
		1	2	3	4	5	6	7	8	9	10	11	12
<u>Estimación no Robusta.</u>	Primera Diferencia Regular, Primera Diferencia Estacional de Incremento Relativo.	-0.46	-0.03	0.22	-0.31	0.30	-0.06	-0.23	0.27	-0.15	0.02	0.24	-0.42
		13	14	15	16	17	18	19	20	21	22	23	24
		0.12	0.09	-0.14	0.10	-0.14	0.02	0.01	0.00	-0.15	0.09	-0.09	0.08
<u>Primera Diferencia Estacional, Primera Diferencia Regular de Incremento Relativo.</u>		R E Z A G O S											
		1	2	3	4	5	6	7	8	9	10	11	12
		-0.22	0.01	0.05	-0.07	0.10	0.04	-0.10	0.11	-0.04	0.03	0.04	-0.25
		R E Z A G O S											
		13	14	15	16	17	18	19	20	21	22	23	24
		-0.03	0.13	-0.02	0.05	0.00	-0.18	0.20	-0.02	-0.11	-0.01	0.04	-0.02

to para el primer como para el segundo modelo indica que la hipótesis que los residuos son ruido blanco no puede rechazarse con un nivel de significación de 25%.

La serie de dinero muestra estacionaridad luego de una diferencia regular de primer orden y de una diferencia estacional también de primer orden sobre el incremento relativo. El correlograma de la Tabla 2 muestra, tanto para estimación robusta como no robusta, autocorrelaciones altas a rezagos 1 y 12 indicando la posibilidad de identificar un modelo multiplicativo con un factor estacional. La Tabla 3 muestra el modelo seleccionado con un parámetro promedio móvil regular y un parámetro promedio móvil estacional. El estadístico Q indica que la hipótesis que los residuos son ruido blanco no puede rechazarse a un nivel de significación de 25%. Este mismo modelo no ha podido estimarse por el momento en forma robusta puesto que no están terminados aun los programas robustos que incluyen modelos multiplicativos estacionales. De cualquier manera ha sido posible computar el estadístico robusto Q^* para este modelo que indica que no puede rechazarse la hipótesis de que los residuos (estimados en forma robusta) son ruido blanco.

Hasta aquí se ha mostrado el efecto del método robusto en la identificación y estimación de modelos ARIMA. A continuación se ilustrará la importancia que tienen los métodos robustos en el análisis de causalidad. Obsérvese en la Tabla 4 que la estimación no robusta de las correlaciones cruzadas indican que debería rechazarse la hipótesis de que dinero causa precios a un nivel de significación de 10% mientras que no es posible rechazar la hipótesis que precios

Tabla N° 3

Estimaciones de Modelos ARIMA para Precios y Dinero.

VARIABLE	PARAMETROS AUTORREGRESIVOS	PARAMETROS PROMEDIO MOVIL	ESTADISTICO Q. (GRADOS DE LIBERTAD)
<u>Estimación no Robusta.</u> Primera Diferencia Incremento Relativo de Precios.	1-0.348B ⁹	1-0.249B-0.395B ² -0.221B ³	Q = 16.9 (16)
Primera Diferencia Estacional, Primera Diferencia Regular de Incremento Rela- tivo de <u>Dinero.</u>		(1-0.571B)(1-0.524B ¹²)	17.2 (18) Q* = 15.4 (22)
<u>Estimación Robusta.</u> Primera Diferencia Incremento Relativo de Precios.		1-0.495B-0.140B ²	Q* = 18.15 (22)

Nota: Q* es el estadístico computado con métodos robustos (ver Apéndice).

Tabla N° 4

Correlaciones Cruzadas entre Precios y Dinero

VARIABLE	R E Z A G O S											
	1	2	3	4	5	6	7	8	9	10	11	12
<u>Estimación no Robusta</u>												
Dinero vs. Precios	0.02	0.20	0.35	-0.08	-0.05	0.03	-0.15	0.08	-0.02	-0.06	0.19	-0.25
Precios vs. Dinero	0.13	0.03	0.17	0.01	0.08	0.25	-0.09	0.09	-0.03	-0.20	0.05	-0.15
<u>Estimación Robusta.</u>												
Dinero vs. Precios	-0.03	0.17	0.12	0.06	-0.02	-0.05	0.00	0.11	-0.04	0.04	0.21	-0.02
Precios vs. Dinero	0.04	-0.10	-0.09	-0.07	-0.04	0.19	-0.12	0.19	-0.10	-0.25	-0.03	-0.19

no causa a dinero a un nivel de significación de 1%.

Si se toma en cuenta ahora las estimaciones robustas para correlaciones cruzadas el panorama luce exactamente al revés: se puede rechazar la hipótesis de que dinero no causa precios a un nivel de significación de 2,5%, mientras que no se puede rechazar la hipótesis de que precios no causa a dinero a un nivel de significación de 50%.

Estos tests muestran claramente la importancia de los métodos robustos en el análisis y especificación de modelos económicos que trabajan con series temporales. Un análisis no robusto inducen a rechazar el modelo monetarista simple de las ecuaciones (3) y (11) (ver, a propósito, las conclusiones obtenidas en Fernández (1979)), cuando, en realidad, un análisis robusto no permite rechazar tal modelo.

IV. Resumen y Conclusiones.

El problema de las perturbaciones en series de tiempo de países como Argentina debe ser considerado con mucha cautela puesto que fácilmente puede inducir a conclusiones erróneas. Un método apropiado para manejar este problema es análisis robusto de series de tiempo que permite obtener estimadores, tanto para parámetros de modelos de series de tiempo como para test de ajuste y causalidad, que se comportan en forma estable frente a unas pocas perturbaciones en los errores o en el modelo.

Una ilustración adecuada de las propiedades del método robusto lo constituye el análisis de la series de dinero y precios para Ar-

gentina en el período 1970-1978. Mientras que el método no robusto induce a rechazar la hipótesis de dinero exógeno (dinero causa a precios), exactamente ocurre lo opuesto cuando el análisis se lleva a cabo en términos robustos.

APENDICE

Aspectos Metodológicos.

En este Apéndice se discuten los diferentes procedimientos robustos que se utilizan en el trabajo. De esta manera se analizan los problemas de robustez en:

- 1) Identificación del modelo ARIMA.
- 2) Estimación de los parámetros del modelo.
- 3) Estimación de residuos.
- 4) Tests de bondad de ajuste.
- 5) Test de Causalidad entre 2 series.

1) Métodos de identificación robusta.

Las técnicas de identificación robusta propuestas en Box and Jenkins (1970) se basan en estimaciones de autocorrelaciones parciales. En este trabajo se emplean estos mismos principios pero utilizando estimadores robustos de las autocorrelaciones y autocorrelaciones parciales.

El método que se utilizó para obtener estimadores robustos de las autocorrelaciones está basado en el trabajo de Maronna (1976) donde se proponen M-estimadores para la estimación robusta de la matriz de covarianza de un vector de observaciones independientes.

Dada una serie estacionaria $z_t, t = 1, T$ la autocorrelación de orden i está dada por:

$$\rho(i) = \frac{\gamma(i)}{\gamma(0)},$$

donde $\gamma(i) = E((z_t - \mu)(z_{t+i} - \mu))$ donde $\mu = E(z_t)$. El procedimiento para estimar robustamente $\rho(i)$ que se ha utilizado es el siguiente. Primero se estima μ por medio de el M-estimador $\hat{\mu}$ que se obtiene resolviendo la ecuación:

$$\sum_{t=1}^T \psi((z_t - \mu)/\hat{\sigma}, k),$$

donde, $\hat{\sigma}$ es un estimador robusto de escala de las z_t y está dado por:

$$\hat{\sigma} = \text{mediana } |z_t - M_z| / 0.675,$$

donde M_z es la mediana de las z_t . La constante 0.675 de normalización se elige de manera que $\hat{\sigma}$ converja a la desviación standard cuando las z_i son normales. Para un estudio más detenido de M-estimadores para un parámetro de locación, ver por ejemplo Huber (1964).

La función $\psi(x, k)$ es la función propuesta por Huber (1964), y está dada por:

$$1) \psi(x, k) = \text{sg}(x) (\text{Max}(|x|, k)),$$

donde, $\text{sg}(x)$ denota la función signo. La constante k se eligió igual a 1.345 que es el valor que hace que el procedimiento tenga eficiencia 0.95 en el caso de que las z_t sean normales.

Una vez obtenido $\hat{\mu}$ se centran los valores de z_t , obteniéndose:

$$z_t^* = z_t - \hat{\mu}.$$

Luego se procede a estimar $\gamma(i)$ y $\gamma(0)$ simultáneamente por medio del siguiente procedimiento iterativo. Los estimadores iniciales serán los usuales:

$$\hat{\gamma}_0(0) = \frac{\sum_{t=1}^T z_t^{*2}}{T-1},$$

$$\hat{\gamma}_0(i) = \frac{\sum_{t=1}^{T-i} z_t^* z_{t+i}^*}{T-i-1}.$$

Sea $\hat{\Gamma}_0$ el estimador inicial de la matriz de covarianza de z_t y z_{t+i} , luego:

$$\hat{\Gamma}_0 = \begin{bmatrix} \hat{\gamma}_0(0) & \hat{\gamma}_0(1) \\ \hat{\gamma}_0(1) & \hat{\gamma}_0(0) \end{bmatrix}.$$

Luego se define al peso inicial w_{t0} del par $\tilde{z}_t^* = (z_t^*, z_{t+1}^*)$ por:

$$w_{t0} = \psi(d_i, k) / d_i \beta,$$

donde:

$$d_i = \tilde{z}_t^{*'} \Gamma_0^{-1} \tilde{z}_t^*,$$

donde:

$$\beta = E\psi(X, k),$$

donde X tiene distribución chi-cuadrado con 2 grados de libertad.

Luego se definen nuevos estimadores de $\gamma(0)$ y $\gamma(i)$ por:

$$\hat{\gamma}_1(0) = \frac{\sum_{t=1}^{T-i} w_{t0} z_t^{*2}}{T-i-1},$$

y

$$\hat{\gamma}_1(i) = \frac{\sum_{t=1}^{T-i} w_{t0} z_t^* z_{t+i}^*}{T-i-1}.$$

Luego el procedimiento se itera hasta que en 2 iteraciones sucesivas se obtienen valores que difieren en menos de un valor pre-establecido.

A estos últimos valores los denominaremos $\hat{\gamma}(0)$ y $\hat{\gamma}(1)$.

El estimador de $\rho(i)$ se obtiene haciendo:

$$\hat{\rho}(i) = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)} .$$

Para obtener estimaciones robustas de las autocorrelaciones parciales se utiliza las ecuaciones de Yule-Walker reemplazando las autocorrelaciones por los estimadores robustos expuestos anteriormente.

2) Estimación robusta de un modelo ARIMA (p,q).

Supongamos que $z_1 \dots z_T$ sea una serie de observaciones centradas que responde a un modelo ARIMA (p,q) estacionario. Luego:

$$2) \phi(B)z_t = \theta(B)u_t,$$

donde,

$$3) \phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p,$$

y

$$4) \theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q.$$

B es el operador de rezago definido por:

$$Bz_t = z_{t-1},$$

y u_t es ruido blanco, es decir las u_t son variables independientes con una misma distribución que llamaremos F .

Box and Jenkins (1970) proponen un procedimiento para estimar los parámetros ϕ_1, \dots, ϕ_p y $\theta_1, \dots, \theta_q$ que es asintóticamente equivalente a máxima verosimilitud en el caso de F normal, y por lo tanto asintóticamente eficiente en este caso.

El procedimiento consiste en minimizar:

$$\sum_{t=1}^T E^2(u_t | \underline{\phi}, \underline{\theta}, \underline{z}),$$

donde $E(u_t | \underline{\phi}, \underline{\theta}, \underline{z})$ es la esperanza condicional de u_t dado:

$\underline{z} = (z_1, \dots, z_T)$ cuando los parámetros son $\underline{\phi} = (\phi_1, \dots, \phi_p)$ y $\underline{\theta} = (\theta_1, \dots, \theta_q)$.

A este estimador lo denominaremos Box y Jenkins (BJ). Un procedimiento aproximado pero asintóticamente equivalente está dado por el estimador de mínimos cuadrados (MC).

El MC-estimador se obtiene minimizando:

$$\sum_{t=p+1}^T \hat{u}_t^2,$$

donde:

$$6) \hat{u}_t = z_t - \phi_1 z_{t-1} - \dots - \phi_p z_{t-p} + \theta_1 \hat{u}_{t-1} + \dots + \theta_q \hat{u}_{t-q},$$

$$\text{con: } \hat{u}_p = \hat{u}_{p-1} = \dots = \hat{u}_{p-q+1} = 0.$$

A pesar que los estimadores BJ y MC son asintóticamente eficientes bajo normalidad de los errores, ambos procedimientos son extremadamente sensibles a la presencia de unas pocas observaciones anómalas.

Por ejemplo consideremos el modelo de contaminación aditiva presentado por Denby y Martin (1979) donde:

$$z_t = x_t + v_t,$$

donde x_t sigue un proceso ARIMA (p,q).

$$\phi(B)x_t = \theta(B)u_t,$$

donde $\phi(B)$ y $\theta(B)$ están definidos como en (3) y (4) con u_t ruido blanco gaussiano y donde los v_t son independientes, siendo $v_t = 0$ con probabilidad $(1-\varepsilon)$ y con probabilidad ε , v_t es igual a una variable con colas pesadas. Esto significa que con una frecuencia $1-\varepsilon$ las observaciones pertenecen a un ARIMA (p,q) y con frecuencia ε se observa el proceso ARIMA (p,q) más un error grosero. Denby y Martin (1979) muestran que en caso autoregresivo el procedimiento MC es altamente inestable para el modelo de contaminación aditiva aún para valores de $\varepsilon = .1$, es decir cuando hay 10% de observaciones anómalas.

Para el caso de que haya un operador de promedios móviles la inestabilidad de los procedimientos MC (y de BJ) es aún mayor, ver Yohai (1980). Lo cual se explica debido a que una contaminación en una observación afecta al cálculo de todos los residuos posteriores.

Por ejemplo la recursión en (6) para un modelo MA(1) da la siguiente expresión para $u_t(\theta_1)$.

$$7) \hat{u}_t(\theta_1) = z_t + \theta z_{t-1} + \theta^2 z_{t-2} + \dots + \theta^{t-1} z_1.$$

Luego se ve que por ejemplo si z_i está afectada de un error grosero esto afectará a todos los \hat{u}_t con $t > i$.

Para el caso de modelos autoregresivos puros AR(p) Denby y Martin (1979) proponen como estimadores robustos de ϕ los estimadores del tipo de máxima verosimilitud generalizados (GM-estimadores). Estos estimadores están definidos como valor de ϕ que minimiza:

$$\sum_{t=p+1}^T \rho(z_t - \phi_1 z_{t-1} - \dots - \phi_p z_{t-p}) w((z_{t-1}' \hat{M}^{-1} z_{t-1})^{1/2}),$$

donde $\rho(u)$ es una función, no decreciente con derivada ψ acotada (7) y donde w es una función no negativa y no decreciente tal que $xw(x)$ es acotada. Por otro lado:

$$\underline{z}_t = (z_t, z_{t-1}, \dots, z_{t-p}),$$

y \hat{M} es un estimador de la matriz de covarianza de \underline{z}_t . Los estimadores son estimadores robustos y las funciones ψ y w pueden elegirse de manera que sean al mismo tiempo eficientes para el caso de u_t normal y sin contaminación (por ejemplo 95% con respecto a BJ o MC) y poco sensibles a la presencia de un cierto número de observaciones afectadas de errores groseros.

Por ejemplo se puede elegir ρ tal que $\rho'(u) = \psi(u, k_1)$ es decir una función de la familia de Huber y dada por (1) y $w(x) = \psi(x, k_2)/x$. Las constantes k_1 y k_2 pueden ajustarse para que se cumpla la condición de 95% de eficiencia en el caso normal.

Otra propuesta hecha por Yohai (1980), que es asintóticamente equivalente a los GM-estimadores, consiste en resolver el siguiente sistema de ecuaciones:

$$\sum_{i=0}^{T-j} d_i(\phi) s_{i+j}(\phi) = 0 \quad j = 1, \dots, p,$$

donde $d_i(\phi)$ son los coeficientes del operador.

$$(1 - \phi_1 B - \dots - \phi_q B^q)^{-1}.$$

Es decir:

$$(1 - \phi_1 B - \dots - \phi_q B^q)^{-1} = \sum_{i=0}^{\infty} d_i(\phi) B^i,$$

y donde:

$$s_k(\phi) = \sum_{i=1}^{T-k} \psi(\hat{u}_t(\phi)/\hat{\sigma}, k) \psi(\hat{u}_{t+k}(\phi)/\hat{\sigma}, k),$$

donde $\hat{\sigma}$ es un estimador de escala de los residuos $u_t(\phi)$. Por ejemplo:

$$\hat{\sigma} = \text{mediana } |\hat{u}_t(\phi)| / .675$$

Para el caso de un modelo ARIMA que contiene un operador de promedio móviles, la estimación robusta de los parámetros se complica por el hecho anteriormente mencionado que una observación anormal contamina los residuos calculados de todos los correspondientes a observaciones posteriores. Esto hace que no sea conveniente utilizar estimadores basados en residuos. En este trabajo se utiliza estimadores robustos propuestos por Yoahi (1980) y que están basados en estimadores estudiados por Tuan (1979) y por Godolphin (1977) y (1978). Estos estimadores a su vez están basados en la idea de Walker (1964) de maximizar una aproximación a la función de verosimilitud que esta basada en las correlaciones muestrales.

Se puede mostrar que los BJ-estimadores son asintóticamente equivalentes a los estimadores de Walker (W-estimadores) que se ob-

tienen minimizando:

$$8) \sum_{t=0}^{T-1} h_i(\tilde{\theta}, \tilde{\phi}) r_i,$$

donde $r_0 = 1$ y,

$$r_i = \frac{\sum_{t=1}^{T-i} z_t z_{t+i}}{\sum_{t=1}^T z_t^2} \text{ para } i > 1,$$

y donde:

$$h_i(\tilde{\theta}, \tilde{\phi}) = \sum_{j=0}^{\infty} d_j(\tilde{\theta}, \tilde{\phi}) d_{j+i}(\tilde{\theta}, \tilde{\phi}),$$

y donde los d_j son los coeficientes del operador $\theta(B)^{-1} \phi(B)$, es decir:

$$(1 - \theta B - \dots - \theta_q B^q)^{-1} (1 - \phi_1 B - \dots - \phi_p B^p) = \sum_{j=0}^{\infty} d_j(\tilde{\theta}, \tilde{\phi}) B^j.$$

Yohai (1980) propuso reemplazar en (8) los r_i por los estimadores robustos $\hat{\rho}(i)$ descritos en el Apéndice, 1. Estos estimadores, que llamaremos "estimadores de Walker modificados" (WM-estimadores), serán robustos y podrán hacerse, en el caso normal, con eficiencia tan próxima a 1 como se quiere eligiendo convenientemente la constante k en la construcción de los $\hat{\rho}(i)$. Estos estimadores son los utilizados en este trabajo para estimar modelos de promedio móviles puros.

3) Estimación robusta de los residuos.

Supongamos que se han obtenido estimadores $\hat{\phi}$, $\hat{\theta}$ de los parámetros del modelo ARIMA (p,q) dado por (2), (3) y (4). Luego los es-

timadores de los residuos \hat{u}_t se obtienen usualmente por el sistema recursivo de ecuaciones (6), reemplazando ϕ y θ por $\hat{\phi}$ y $\hat{\theta}$.

Estos residuos pueden ser utilizados, por ejemplo, para testear la bondad de ajuste del modelo utilizado, como se verá en el Apéndice 4, o, para testear causalidad con otra serie, como se verá en el Apéndice 5.

Sin embargo, si el modelo contempla un operador de promedios móviles, aún utilizando estimadores robustos de $\hat{\phi}$ y $\hat{\theta}$, unas pocas observaciones anómalas pueden afectar a un gran número de valores \hat{u}_t calculados por medio de la fórmula (6). Esto ya se vió para el caso de un MA(1) en donde \hat{u}_t estaba dada por (7). Por lo tanto para evitar la propagación de los efectos de las observaciones anómalas sobre una gran cantidad de residuos proponemos modificar (6) por:

$$\hat{u}_t = z_t^{-\phi} z_{t-1} \dots - \phi_p z_{t-p} + \theta_1 \hat{\sigma} \psi(\hat{u}_{t-1}/\hat{\sigma}, k) \dots + \theta_q \hat{\sigma} \psi(\hat{u}_t/\hat{\sigma}, k),$$

donde $\hat{\sigma}$ es un estimador de escala de los residuos y que puede estar dado por ejemplo por:

$$\hat{\sigma} = \text{mediana } |\hat{u}_t| / .675.$$

En este trabajo se ha tomado como $k = 2$ de manera que si los u_t son normales se modifiquen aproximadamente el 5% de ellos.

4) Test de bondad de ajuste del modelo.

Para testear la bondad de ajuste de un modelo, usualmente se utiliza el test de Box y Pierce. Este test se basa en el estadís-

tico:

$$* \quad Q = \sum_{i=1}^N r_{ui}^2,$$

dónde las r_{ui} son las autocorrelaciones empíricas de los residuos estimados mediante el sistema (6), y donde se utilizan BJ o MC-estimadores. Los r_{ui} están definidos por:

$$9) \quad r_{ui} = \frac{\sum_{t=1}^{T-i} \hat{u}_t \hat{u}_{t+i}}{\sum_{t=1}^T \hat{u}_t^2}$$

Box y Pierce (1976) demuestran que Q está distribuido asintóticamente como una distribución chi-cuadrado con $N-p-q$ parámetros. Para robustificar este test se propone en este trabajo utilizar los \hat{u}_t calculados como se explicó en el Apéndice, 3. De todos modos en caso de haber algunas observaciones anómalas esto será reflejado en los \hat{u}_t . Para evitar la influencia de estas se propone reemplazar los \hat{u}_t por $\hat{v}_t = \psi(\hat{u}_t/\hat{\sigma}, k)$ y definir r_{vi} en forma análoga a (9) y luego definir:

$$Q^* = \sum_{i=1}^N r_{vi}^2$$

El estadístico Q^* se deberá también distribuir aproximadamente para T grande como una chi-cuadrado con $N-p-q$ grados de libertad. Luego el test de bondad de ajuste del modelo se hace en base al estadístico Q^* .

5) Tests robustos de causalidad entre 2 series.

Sean 2 series $(x_t, y_t) t = 1, \dots, T$ tales que siguen procesos ARIMA (p_1, d_1, q_1) y ARIMA (p_2, d_2, q_2) respectivamente. Sean u_{1t} y

u_{2t} las innovaciones correspondientes de ambas series. Definamos además para todo k entero $-\infty < k < \infty$,

$$\rho_{u_1, u_2}(k) = E(u_{1,t} u_{2,t+k}) / (E(u_{1,t}^2) E(u_{2,t}^2))^{1/2}.$$

Siguiendo a Pierce (1977) diremos que "x causa y" si $\rho_{u_1, u_2}(k) \neq 0$ para algún $k > 0$ y diremos que "y causa x" si $\rho_{u_1, u_2}(k) \neq 0$ para algún $k < 0$.

Sea H la hipótesis nula que las series $u_{1,t}$ y $u_{2,t}$ son independientes. Luego para testear la hipótesis nula H contra la alternativa "x causa y" Haugh (1976) propone un test basado en:

$$Q_{xy} = T \sum_{k=1}^N r_{u_1, u_2}^2(k),$$

donde,

$$10) r_{u_1, u_2}(k) = \frac{\sum_{t=1}^{T-k} \hat{u}_{1,t} \hat{u}_{2,t+k}}{\left(\left(\sum_{t=1}^T \hat{u}_{1,t}^2 \right) \left(\sum_{t=1}^T \hat{u}_{2,t}^2 \right) \right)^{1/2}},$$

y donde $\hat{u}_{1,t}$ y $\hat{u}_{2,t}$ son residuos estimados de las 2 series luego de estimar los parámetros de los modelos por medio de BJ y MC-estimadores.

Haugh (1976) demuestra que el estadístico Q_{xy} tiene bajo H distribución asintótica chi-cuadrado con N grados de libertad.

Análogamente para testear H contra "y causa x" se puede usar:

$$Q_{yx} = T \sum_{k=-N}^{-1} r_{u_1, u_2}^2(k).$$

Para robustificar estos tests se pueden utilizar $\hat{u}_{1,t}$ y $\hat{u}_{2,t}$

tal como se proponen en el Apéndice 3., y luego similarmente a lo visto en el Apéndice 4. estos valores se robustifican reemplazándolos por:

$$\hat{v}_{i,t} = \psi(\hat{u}_{i,t}/\hat{\sigma}_i, k) \quad i = 1, 2,$$

donde:

$$\hat{\sigma}_i = \text{mediana} (|\hat{u}_{i,t}|) / .675 \quad i = 1, 2.$$

Luego en forma similar a (10) se define $r_{v_1, v_2}(k)$ y los estadísticos:

$$Q_{xy}^* = T \sum_{k=1}^N r_{v_1, v_2}^2(k),$$

$$Q_{yx}^* = T \sum_{k=-N}^{-1} r_{v_1, v_2}^2(k).$$

Tanto Q_{xy}^* y Q_{yx}^* tendrán asintóticamente la misma distribución chi-cuadrado con N grados de libertad.

REFERENCIAS

- Box, G.E.P. and Jenkins, G.M. (1970). Time Series Analysis. Holden Day.
- Box, G.E.P. and Pierce, D.A. (1970). Distribution of residual autocorrelations in autoregressive integrated moving average time series. Journal American Statistical Association, 64.
- Denby, M. and Martin, R.D. (1979). Robust estimation of the first-order autoregressive parameter. Journal American Statistical Association, 74.
- Fernández, Roque B. (1979). Dinero y Precios: su Interrelación en el corto plazo, Documentos de Trabajo, CEMA, N° 7, Noviembre.
- Godolphin, E.J. (1977). A direct representation for maximum likelihood estimation of a Gaussian moving average process. Biometrika 65.
- Godolphin, E.J. (1978). Modified maximum likelihood estimation of Gaussian moving average process using a pseudo quadratic convergence criterion. Biometrika 65.
- Haugh, L.D. (1976). Checking the independence of two covariance stationary time series: A univariate residual cross correlation approach. Journal American Statistical Association, 71.
- Huber, P. (1964). Robust estimation of a location parameter. Ann. of Math. Stat. 35.
- Maronna, R. (1976). Robust M-estimates of multivariate location and scatter. Ann. of Statistics, 4.
- Pierce, D.A. (1977). Relationships -and the lack there of- between economic time series with special reference to money and interest rates. Journal American Statistical Association, 72.

- Tuan, P.D. (1979). The estimation of parameters for autoregressive-moving average models from sample autocovariances. *Biometrika*, 66.
- Walker, A.M. (1964). Asymptotic properties of least squares of parameter of the spectrum of a Stationary non-deterministic time series. *J.Aust. Math. Sec.*, 4.
- Yohai, Víctor J. (1980). Robust Estimation in Autoregressive Moving Average Models. Trabajo presentado en Seminario de Series de Tiempo, Banco Central-C.E.M.A., Julio 1980.