

Inversión y PNB: Relaciones de causalidad

Alvaro Montenegro García

Este artículo investiga la relación de causalidad entre la inversión y el PNB por medio de la prueba de causalidad de Granger, utilizando para ello datos norteamericanos de la posguerra. Un ejercicio de esta naturaleza es importante por el énfasis que con frecuencia se asigna a la inversión, o a diferentes componentes de la misma, dentro del diseño de la política macroeconómica.

Desde el punto de vista filosófico, la definición de causalidad generalmente aceptada es aquella que se refiere a "predicción de acuerdo a un conjunto de leyes". Quiere decir esto que si contamos con un conjunto de leyes, que pueden ser leyes físicas o, en nuestro caso, leyes económicas, y podemos hacer predicciones de alguna variable con base en esas leyes y en otras variables que entran en la estructura de dichas leyes, entonces también podemos establecer relaciones causales inequívocas entre las variables.

El problema es que en la economía tenemos pocas leyes. Tal vez todos estaríamos de acuerdo en mencionar la Ley de la Oferta y la Demanda; pero, por ejemplo, si adicionalmente mencionamos la Ley de Say empezamos a perder consenso. Y perderemos aún mas consenso si insistimos en mencionar leyes como las de Okun, Verdoon o Murphy.

En econometría, la prueba de Granger (1969) ciertamente incorpora el concepto de "predicción" de la definición filosófica pero sustituye las "leyes" por la correlación estadística entre variables. Sin embargo, goza de amplia aceptación y, junto con la prueba de Sims (1972), son las de mayor utilización en estudios econométricos de causalidad.

I. Causalidad en el sentido de Granger

Se dice que la variable X causa a Z en el sentido de Granger cuando, en presencia del pasado de Z , el pasado de X ayuda a explicar Z . Esta definición se torna operacional al interpretarla dentro de un contexto lineal pudiendo así implementarla en términos econométricos más familiares: corremos una regresión con Z como variable dependiente y con rezagos de Z y X como variables explicativas; y si encontramos que alguno de los rezagos de X juega un papel significativo en la explicación de Z entonces decimos que X causa a Z en el sentido de Granger. Para determinar si la información contenida en los rezagos de X es significativa se utiliza una prueba F . Igualmente, diremos que Z causa a X si, en presencia del pasado de X , el pasado de Z ayuda a explicar X . Para ilustrar estos conceptos consideremos las siguientes ecuaciones de regresión:

$$Z_t = a + \sum_{i=1}^{\infty} b_i Z_{t-i} + \sum_{j=1}^{\infty} c_j X_{t-j} + e_t \quad (1)$$

$$X_t = d + \sum_{i=1}^{\infty} f_i Z_{t-i} + \sum_{j=1}^{\infty} g_j X_{t-j} + u_t \quad (2)$$

donde a y d son constantes, b_i , c_j , f_i y g_j son los coeficientes y e_t y u_t son los residuos. Diremos, por ejemplo, que hay causalidad unidireccional de X hacia Z si alguno de los coeficientes c_j es significativamente diferente de cero mientras que los coeficientes f_i son cero. Ninguna de las dos variables es causal si tanto c_j como f_i son cero; y si algunos coeficientes de c_j y f_i son diferentes de cero diremos que X y Z son mutuamente causales. Una exposición más rigurosa de estas definiciones de causalidad puede consultarse en el artículo de Granger.

Sims (1972) afirma que la causalidad unidireccional en el sentido de Granger es equivalente a la exogeneidad econométrica, de manera que este tipo de pruebas resulta de gran ayuda en la especificación de modelos. Engle, Hendry y Richard (1983) también relacionan esta definición de causalidad con la exogeneidad. Pierce (1975) sugiere que el concepto de indicador líder se utilice para denotar una variable que cause a otra en el sentido de Granger.

Algunas de las críticas formuladas a la prueba de Granger (Judge 1985, p. 668) tienen que ver con el criterio de predicción utilizado para determinar causalidad en lugar del más comprensivo concepto de causa y efecto; con la escogencia y manera de aplicar el error cuadrado medio (varianza) como indicador para juzgar la bondad del predictor; con el hecho de limitarse a predictores lineales (la regresión) excluyendo predictores no lineales, y con el hecho de limitar la información disponible al pasado y presente de unas pocas variables (Clavijo 1989), X y Z en este caso.

En el texto que sigue el término causalidad no tiene connotación filosófica alguna y cuando digamos, por ejemplo, que X no causa a Z , entenderemos simplemente que las predicciones de Z no pueden ser mejoradas conociendo el pasado de X .

II. Algunas consideraciones prácticas

Es importante trabajar con series largas pues de otra forma el número de rezagos y de parámetros a estimar reduce considerablemente los grados de libertad. Además, por

contener variables dependientes rezagadas, las regresiones y pruebas estadísticas indicadas en (1) y (2) sólo son válidas asintóticamente (Johnston 1974, p. 362).

La periodicidad de las series de tiempo debe ser la adecuada para detectar la relación temporal entre las variables. Por ejemplo, no deben emplearse series anuales si se sospechan efectos causales trimestrales o mensuales.

La prueba de Granger está diseñada para series de tiempo estacionarias; es decir, para series que no tengan ni tendencia, ni ciclos visibles, ni sean explosivas o implosivas; sino para series cuyo aspecto visual no cambie a través del tiempo. En términos más formales, son series estacionarias aquellas cuyos momentos (media, varianza, etc.) sean invariables en el tiempo. Usualmente basta con que la media y la varianza sean constantes (lo cual se conoce como estacionariedad de orden dos).

Algunos de los métodos más comunes para convertir una serie no estacionaria en estacionaria incluyen tomar variaciones absolutas, tomar logaritmos, variaciones de logaritmos, variaciones porcentuales y la desestacionalización. Pierce y Haugh (1977) indican que estas transformaciones lineales y que algunas no lineales, como la logarítmica, no alteran las relaciones de causalidad entre las variables.

Usualmente se supone que los errores no están correlacionados entre sí, de manera que $E[e_t u_s] = 0$ para todo t y s . De otra forma debemos recurrir a métodos de estimación para ecuaciones simultáneas.

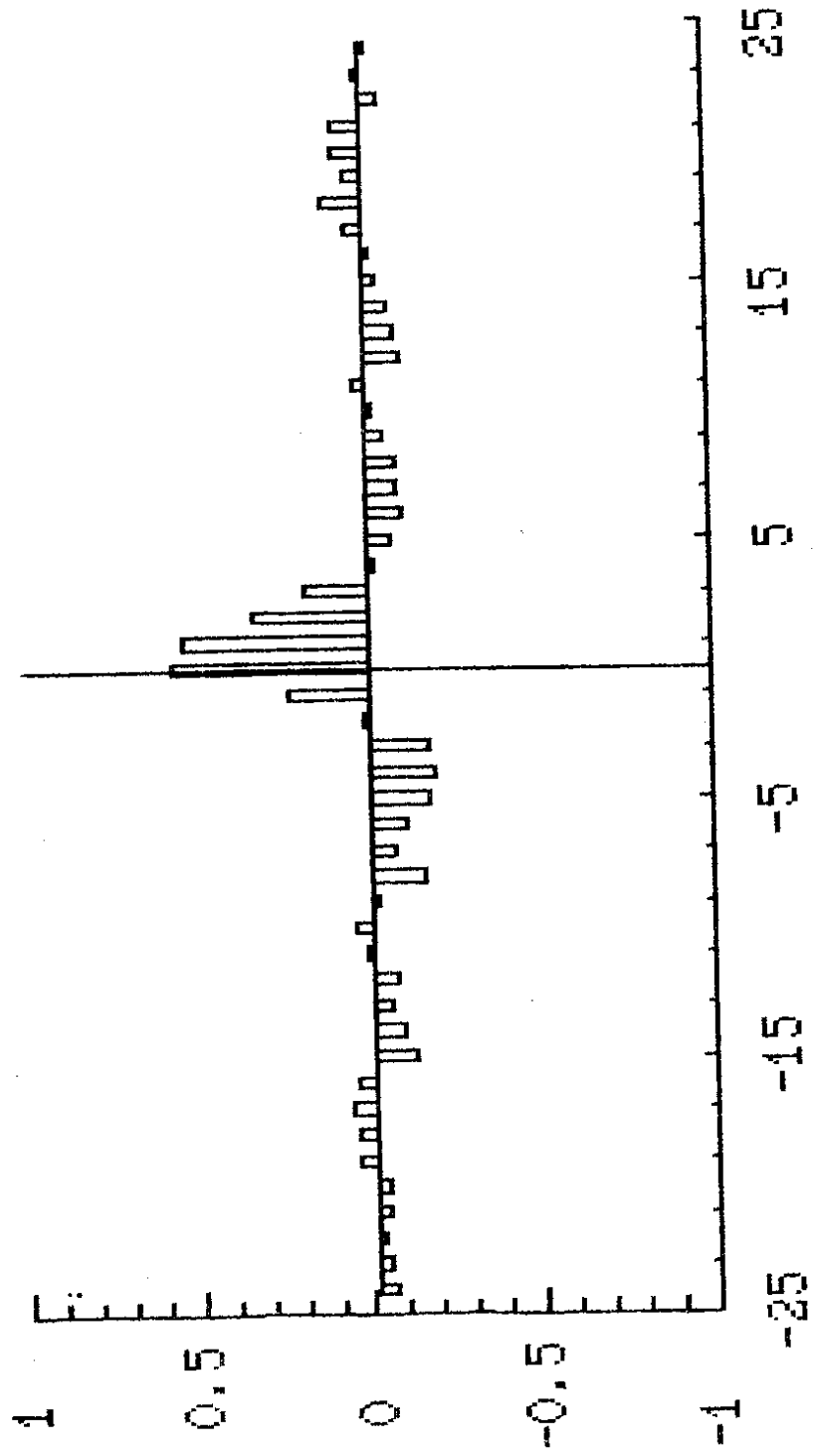
El número de rezagos indicados en las ecuaciones (1) y (2) es infinito, sin embargo, por consideraciones prácticas éste tendrá que ser finito y menor que la longitud de las series. Es usual que la longitud de los rezagos se fije o de manera algo arbitraria, o con base en criterios tipo Akaike como el del error final de predicción (ver Thornton y Batten 1985, o Hall y Noble 1987), o con base en la información contenida en la función de correlación cruzada. La omisión de rezagos importantes de la variable dependiente puede generar autocorrelación en los residuos y sesgar los coeficientes de la variable independiente, sesgando, en consecuencia, la prueba F .

III. La inversión y el PNB

Se estudiarán las dos siguientes series estadounidenses: (1) el cambio porcentual en el PNB, que llamaremos Y , y (2) el cambio porcentual en la inversión bruta privada doméstica no residencial (planta y equipo), que llamaremos I . Las series originales, desestacionalizadas, trimestrales y en precios constantes de 1982, fueron tomadas de la publicación National Income and Product Accounts. El período cubierto por las series porcentuales va desde 1947/2 hasta 1987/4, para un total de 163 observaciones.

La escogencia de la longitud de los rezagos se hizo con base en la función de correlación cruzada mostrada en el gráfico adjunto. Este muestra la correlación entre la serie Y y la serie I , rezagando la serie Y con respecto a la serie I ; y sugiere que 6 rezagos son suficientes para ambas variables. Las ecuaciones (1) y (2) quedan como sigue:

FUNCION DE CORRELACION CRUZADA
PNB% vs. INVERSION NO RESIDENCIAL%



Rezago del PNB con respecto a INVERSION

$$I_t = a + \sum_{i=1}^6 b_i I_{t-i} + \sum_{j=1}^6 c_j Y_{t-j} + e_t \quad (3)$$

$$Y_t = d + \sum_{i=1}^6 f_i I_{t-i} + \sum_{j=1}^6 g_j Y_{t-j} + u_t \quad (4)$$

La prueba F se realiza bajo la hipótesis nula de que la variable independiente no causa la dependiente. Para ello es necesario comparar la variación no explicada (suma de residuos elevados al cuadrado) de la regresión (3) con la variación no explicada cuando hacemos $c_j=0$, para todo j , en la misma ecuación (3). De manera similar para la ecuación (4) corremos una regresión sin restricciones y otra fijando $f_i=0$, para todo i . En otras palabras, corremos las siguientes dos parejas de regresiones:

$$I_t = a + \sum_{i=1}^6 b_i I_{t-i} + \sum_{j=1}^6 c_j Y_{t-j} + e_t \quad (5)$$

$$I_t = a' + \sum_{i=1}^6 b_i' I_{t-i} + e_t' \quad (6)$$

cuyos resultados se muestran en el Cuadro 1, y

$$Y_t = d + \sum_{i=1}^6 f_i I_{t-i} + \sum_{j=1}^6 g_j Y_{t-j} + u_t \quad (7)$$

$$Y_t = d' + \sum_{j=1}^6 g_j' Y_{t-j} + u_t' \quad (8)$$

cuyos resultados se muestran en el Cuadro 2.

La prueba F requiere, para su validez, que los residuos sean ruido blanco (que no exista autocorrelación). Para probar esta independencia lineal se han diseñado varias pruebas, una de las cuales es el estadístico Q, donde

$$Q = n(n+2) \sum_{\tau=1}^p (n-\tau)^{-1} r_{\tau} \quad (9)$$

el cual, bajo la hipótesis nula de que la serie de residuos es ruido blanco, tiene una distribución chi-cuadrado con p grados de libertad. La expresión (9) corresponde a una versión para muestras moderadas (ver Feige y Pearce 1979) y será la utilizada aquí. El número de observaciones n en la serie de residuos es de 157 pues de las 163 originales se pierden 6 al tomar rezagos; el número de autocorrelaciones consideradas p se fijó en 24; y r_{τ} es el coeficiente de autocorrelación de la serie de residuos para el rezago τ , calculado según

$$r_{\tau} = (n\sigma_e^2)^{-1} \sum_{t=\tau}^n e_t e_{t-\tau} \quad (10)$$

donde σ_e^2 es la varianza de e_t (ver Montenegro 1989). Un valor significativamente alto de Q llevaría a rechazar la hipótesis nula de que la serie considerada es ruido blanco, lo cual, a su vez, nos llevaría a sospechar de la validez de la prueba F.

IV. Resultados

Los valores de Q para los residuos de las regresiones, mostradas en los Cuadros 1 y 2, no son estadísticamente significativos al 5% (ver tabla de distribución chi-cuadrado con 24 grados de libertad), indicando la ausencia de correlación serial.

Cuadro 1

Variable dependiente: I_t		# de observaciones: 157	
Variable independiente	coeficiente	estadístico t	confiabilidad
Constante	-.31	-.82	.41
I_{t-1}	.15	1.61	.11
I_{t-2}	-.10	-1.03	.31
I_{t-3}	.06	.63	.53
I_{t-4}	-.02	-.15	.88
I_{t-5}	-.15	-1.53	.13
I_{t-6}	.09	.99	.32
Y_{t-1}	1.04	4.44	.00
Y_{t-2}	.43	1.73	.09
Y_{t-3}	.05	.18	.86
Y_{t-4}	-.21	-.85	.40
Y_{t-5}	.17	.68	.50
Y_{t-6}	-.04	-.17	.86
$R^2 = .3358$ variación explicada = 475.583 $Q = 23.6$ variación no explicada = 940.46			

Variable dependiente: I_t		# de observaciones: 157	
Variable independiente	coeficiente	estadístico t	confiabilidad
Constante	.66	2.66	.01
I_{t-1}	.41	5.06	.00
I_{t-2}	-.02	-.27	.79
I_{t-3}	.06	.65	.52
I_{t-4}	-.10	-1.15	.25
I_{t-5}	-.17	-1.93	.06
I_{t-6}	.08	.91	.36
$R^2 = .2127$ variación explicada = 301.171 $Q = 20.0$ variación no explicada = 1114.82			

Cuadro 2

Variable dependiente: Y_t		# de observaciones: 157	
Variable independiente	coeficiente	estadístico t	confiabilidad
Constante	.42	2.81	.01
Y_{t-1}	.30	3.16	.00
Y_{t-2}	.23	2.32	.02
Y_{t-3}	-.00	-.00	1.00
Y_{t-4}	-.03	-.26	.80
Y_{t-5}	.06	.56	.58
Y_{t-6}	.11	1.06	.29
I_{t-1}	-.00	-.01	1.00
I_{t-2}	-.042	-1.12	.27
I_{t-3}	-.06	-1.57	.12
I_{t-4}	-.03	-.72	.46
I_{t-5}	-.03	-.87	.39
I_{t-6}	-.01	-.41	.68
$R^2 = .2231$ variación explicada = 43.6407 $Q = 13.0$ variación no explicada = 151.948			

Variable dependiente: Y_t		# de observaciones: 157	
Variable independiente	coeficiente	estadístico t	confiabilidad
Constante	.55	3.98	.00
Y_{t-1}	.33	4.06	.00
Y_{t-2}	.20	2.28	.02
Y_{t-3}	-.10	-1.15	.25
Y_{t-4}	-.12	-1.34	.18
Y_{t-5}	-.03	-.36	.72
Y_{t-6}	.04	.50	.62
$R^2 = .1871$ variación explicada = 36.5927 $Q = 12.0$ variación no explicada = 158.996			

El estadístico F, que mide la diferencia en poder explicativo entre las dos regresiones (5) y (6) mostradas en el cuadro 1, se calcula así:

$$F = (1114.82 - 940.46) (144) / (940.46) (6) = 4.449$$

donde los grados de libertad del numerador son 6, igual al número de restricciones; y de 144 para el denominador, igual a 157 observaciones menos 13 parámetros estimados (6 rezagos de I, 6 rezagos de Y y la constante).

El F calculado de 4.449 es significativo al .04% de confiabilidad, indicando que Y causa I en el sentido de Granger; esto es, que el PNB causa la inversión no residencial.

Los resultados para las regresiones (7) y (8) se muestran en el cuadro 2. La prueba F arroja el siguiente cálculo:

$$F = (158.996 - 151.948) (144) / (151.948) (6) = 1.113$$

el cual no es significativo (ver tabla de distribución F para 6 grados de libertad en el numerador y 144 en el denominador). Este resultado implica que la inversión no residencial no causa el PNB.

V. Otros resultados

Se repitió el mismo ejercicio anterior utilizando la inversión residencial en lugar de la no residencial. Los resultados (no mostrados aquí) indican que estos dos tipos de inversión tienen un comportamiento de causalidad opuesto con respecto al PNB; que mientras el PNB causa la inversión no residencial, la inversión residencial causa el PNB. Sin embargo, debemos advertir que la inferencia con respecto a la inversión residencial es un poco menos significativa en términos estadísticos que en el caso anterior, como se puede apreciar comparando la confiabilidad de los valores F estimados.

La prueba F para las regresiones del tipo (5) y (6) utilizando la inversión residencial es de 2.135, significativo al 5.3%. Para las ecuaciones de tipo (7) y (8) el valor de F es de 4.316, significativo al .05%.

Si en lugar de 6 rezagos utilizamos combinaciones de 4 y 8 rezagos, obtenemos resultados (no mostrados aquí) cualitativamente similares tanto para la inversión no residencial y el PNB como para la inversión residencial y el PNB.

VI. Inversión y PIB en Colombia

La realización de un estudio de causalidad similar para Colombia enfrenta varias dificultades. Tal vez la principal es la limitación en el número de observaciones, 22 datos anuales porcentuales en el período 65-87, los cuales se reducen a 15 grados de libertad al efectuar regresiones con solo dos rezagos. Así, las inferencias de la prueba de Granger no resultan muy confiables. Relativo a lo anterior, datos anuales pueden esconder

relaciones importantes de causalidad de periodicidad menor, por ejemplo, trimestrales. Otra dificultad radica en el desfase de la actividad edificadora con respecto a las licencias de construcción utilizadas en la elaboración de las cuentas nacionales.

VII. Conclusiones

El resultado del ejercicio de causalidad de Granger presentado aquí sugiere que el PNB causa (ayuda a predecir) la inversión privada no residencial (planta y equipo) y no al contrario; esto es, que al considerar estas dos variables, el PNB tiende a ser exógeno y la inversión no residencial endógena o, en términos econométricos de regresión, que la inversión no residencial se modela mejor como variable dependiente y el PNB como variable independiente.

En contraste, la prueba indica que la inversión residencial causa (ayuda a predecir) el PNB; sin embargo, también hay evidencia, aunque menos fuerte, de causalidad del PNB hacia la inversión residencial.

De las dos, la residencial y la no residencial, la inversión residencial parece ofrecer mayores ventajas para el diseño de política macroeconómica.

Bibliografía

- Clavijo, Sergio. 1989. "Macroeconometría de una economía pequeña y abierta usando análisis de vectores autorregresivos". *Revista Desarrollo y Sociedad*. No. 23. Marzo.
- Engle R., Hendry D. y Richard J., 1983. Exogeneity. *Econometrica*. Vol. 51. No. 2, 277-304.
- Feige, Edgar y Pearce, Douglas. 1979. "The casual causal relationship between money and income: some caveats for time series analysis". *Review of Economics and Statistics*. Vol. 61. No. 4. Noviembre.
- Granger, Clive, 1969. "Investigating causal relations by econometric models and cross-spectral methods". *Econometrica*. Vol 37. No. 3. Julio.
- Hall, Thomas y Noble, Nicholas. 1987. "Velocity and the variability of money growth: evidence from Granger causality tests". *Journal of Money, Credit and Banking*. Vol 19. No. 1. Febrero.
- Johnston, John. 1984. *Econometric Methods*. McGraw-Hill.
- Judge, G., et. al. 1985. *The Theory and Practice of Econometrics*. Wiley.
- Montenegro, Alvaro. 1989. "La función de autocorrelación y su empleo en el análisis de series de tiempo". *Revista Desarrollo y Sociedad*. No. 23. Marzo.
- Pierce, David. 1975. "Forecasting in dynamic models with stochastic regressors". *Journal of Econometrics*. No. 3. Noviembre.
- Pierce, David y Haugh, Larry. 1977. "Causality in temporal systems". *Journal of Econometrics*. No. 5. Mayo.
- Sims, Christopher. 1972. "Money, income and causality". *American Economic Review*. No. 62.
- Thornton, Daniel, y Batten, Dallas. 1985. "Lag-length selection and test of Granger causality between money and income". *Journal of Money, Credit and Banking*. Vol 17. No. 2. Mayo.