An Expository Note on Sims's Formula Describing Discrete Time Approximations to Continuous Time Distributed Lags

by

Thomas Sargent

January 1976

This purely pedagogical note derives a very important formula due to Sims without resorting to the frequency domain calculations that he used to derive it. Most economists are better able to follow calculations in the time domain than in the frequency domain. The time domain derivation advanced here makes apparent the interpretation of Sims's formula as a version of Theil's specification error theorem.

Sims considered the model

(1) $y(t) = \int_{-\infty}^{\infty} b(s)\, x(t-s)ds + u(t)$

where b(t) is an absolutely integrable function,[1/] and where y(s), x(s), and u(s) are continuous time covariance stationary stochastic processes with means of zero and finite variances. The disturbance process u(s) is orthogonal to the x(s) process, that is

(2) $E[u(s)\, x(t)] = 0$ for all real t, s.

The specification (2) identifies the convolution $\int_{-\infty}^{\infty} b(s)\, x(t-s)ds$, which is denoted $b * x(t)$, as the projection of y(t) on the entire x(s) process, $s \varepsilon (-\infty, \infty)$.[2/]

Sims studied the situation where y(t) and x(t) are only observed at the integers. The data thus consist of the sequences of random variables

$$X(n) = x(n)$$
$$Y(n) = y(n) \quad , \quad n = \ldots, -2, -1, 0, 1, 2, \ldots.$$

Here X(n) and Y(n) are discrete time stochastic processes. Consider the least squares distributed lag regression (i.e., the projection) of Y(n) on past, present, and future X(r)'s:

(3) $$Y(n) = \sum_{s=-\infty}^{\infty} B(s)\ X(n-s) + U(n)$$

where U(n) is the least squares disturbance and where

(4) $$E[U(n)\ X(r)] = 0 \qquad \text{for all integer } n,\ s.$$

Sims was interested in studying the relation between B(s) and b(t) and in investigating the conditions under which B(s) well or poorly represents b(t) sampled at the integers.

To obtain Sims's formula, we begin by recalling that the orthogonality condition (4) uniquely determines the B(s)'s.[3/] Substitute for U(n) from (3) into (4) to obtain

(5) $$E[(Y(n) - \sum_{s=-\infty}^{\infty} B(s)\ X(n-s))\ X(r)] = 0 \qquad \text{for all integer } n,\ r,\ s.$$

or

(6) $$E[Y(n)\ X(r)] - \sum_{x=-\infty}^{\infty} B(s)\ E[X(n-s)\ X(r)] = 0.$$

These are the least squares "normal equations." We write (6) as

$$R_{YX}(n-r) - \sum_{s=-\infty}^{\infty} B(s)\ R_X(n-r-s) = 0$$

or

(7) $$R_{YX}(\tau) - \sum_{s=-\infty}^{\infty} B(s)\ R_X(\tau-s) = 0$$

where we define the covariance sequences

$$R_{YX}(\tau) = E[Y(t)\ X(t-\tau)] \qquad \text{for all } t, \text{ integer } \tau$$

(8)

$$R_X(\tau) = E[X(t+\tau)\ X(t)] \qquad \text{for all } t, \text{ integer } \tau.$$

R_{YX} and R_X are independent of t by virtue of the covariance stationarity of Y and X. It is also convenient to define here the covariance functions

$$R_{yx}(\tau) = E[y(t)\ x(t-\tau)] \qquad \text{for all real } t, \text{ real } \tau$$

(9)

$$R_x(\tau) = E[x(t)\ x(t+\tau)] \qquad \text{for all real } t, \text{ real } \tau.$$

Clearly, the covariance sequences R_{YX} and R_X correspond to the covariance functions R_{yx} and R_x, respectively, sampled at the integers.

Now from (1) we have that

$$R_{YX}(\tau) = E[Y(t)\ X(t-\tau)] = E[(\int_{-\infty}^{\infty} b(s)\ x(t-s)ds + u(t)).\ X(t-\tau)]$$

$$= \int_{-\infty}^{\infty} b(s)\ E(x(t-s)\ X(t-\tau))ds + E\ u(t)\ X(t-\tau)$$

which, applying (2) and (9), equals

$$\int_{-\infty}^{\infty} b(s)\ R_x(\tau-s)ds$$

so that

$$R_{YX}(\tau) = b * R_x(\tau) \qquad \text{for integer values of } \tau.$$

So equation (7) becomes

$$(8) \qquad b * R_x(\tau) = \sum_{s=-\infty}^{\infty} B(s)\ R_X(\tau-s).$$

Now let $R_X^{-1}(s)$ denote the sequence which is the inverse under convolution of the sequence $R_X(s)$. This inverse is defined by[4/]

$$\sum_{j=-\infty}^{\infty} R_X^{-1}(j)\, R_X(n-j) = \begin{cases} 1 & \text{for } n = 0 \\ 0 & \text{for } n \neq 0, \quad n \text{ an integer.} \end{cases}$$

Convoluting the left side of (8) with $R_X^{-1}(j)$ gives

$$\sum_{j=-\infty}^{\infty} R_X^{-1}(j)[b * R_x(\tau-j)] = \sum_{j=-\infty}^{\infty} R_X^{-1}(j) \int_{-\infty}^{\infty} b(s)\, R_x(\tau-j-s)ds$$

(9)

$$= \int_{-\infty}^{\infty} b(s)\Big(\sum_{\tau=-\infty}^{\infty} R_X^{-1}(j)\, R_x(\tau-j-s)\Big)ds.$$

Notice that

$$\sum_{\tau=-\infty}^{\infty} R_X^{-1}(j)\, R_x(\tau-j) = R_X^{-1} * R_x(\tau),$$

which is the convolution of the sequence $R_X^{-1}(j)$ with the function $R_x(t)$.5/ Sims calls this convolution $r_x(\tau) = R_X^{-1} * R_x(\tau)$. Notice that it is defined for all real $\tau - s$. Thus, (9) can be written

$$\int_{-\infty}^{\infty} b(s)\, (R_X^{-1} * R_x(\tau-s))ds$$

$$= b * R_X^{-1} * R_x(\tau).$$

While this convolution exists for all real τ, we are interested in its values only at integer τ (refer again to (8)).

Now convolute the right side of equation (8) with $R_X^{-1}(j)$ to get

$$R_X^{-1} * \sum_{s=-\infty}^{\infty} B(s)\, R_X(\tau-s) = \sum_{j=-\infty}^{\infty} R_X^{-1}(j) \sum_{s=-\infty}^{\infty} B(s)\, R_X(\tau-s-j)$$

$$= \sum_{s=-\infty}^{\infty} B(s) \sum_{j=-\infty}^{\infty} R_X^{-1}(j)\, R_X(\tau-j-s)$$

$$= B(\tau).$$

Combining the results of convoluting the left and right sides of (8) with R_X^{-1} gives

(10) $$B(\tau) = b * R_X^{-1} * R_x(\tau)$$

or more explicitly

$$B(\tau) = \int_{-\infty}^{\infty} b(s) \left(\sum_{j=-\infty}^{\infty} R_X^{-1}(j)\, R_x(\tau-j-s)\right) ds.$$

Equation (10) is Sims's formula. It states that $B(\tau)$ is formed by weighting $b(s)$ by $R_X^{-1} * R_x(\tau-s)$ and then integrating over all real s's. The weighting function $R_X^{-1} * R_x(\tau-s)$ clearly depends on the stochastic structure of the x-process. Sims's paper contains a variety of interesting and useful results about the shape of the weighting function for various classes of x-processes.

Relation to Theil's Specification Theorem

Let the least squares projection of a random variable z on a 1 x k vector of random variables Z be $Z\alpha$ where α is the k x 1 vector of regression coefficients. Partition Z as $Z = (Z_1\ Z_2)$ where Z_1 is $(1xk_1)$ and Z_2 is $(1xk_2)$ with $k_1 + k_2 = k$. Theil's specification theorem states that the projection of z on Z_1 is $Z_1\xi$ where ξ is (k_1x1) and

$$\xi = \underset{k_1xk_1}{\Gamma}\ \underset{kx1}{\alpha}$$

where the projection of the i^{th} element of Z on Z_1 is $\sum_{j=1}^{k_1} \Gamma_{ji} Z_j$; Γ_{ji} is the partial regression coefficient of the i^{th} dependent variable on the j^{th} Z. The preceding equation can be written

(11) $$\xi_i = \sum_{i=1}^{k} \Gamma_{ij}\ \alpha_j\ ,$$

which says that the coefficient on Z_i in the projection of z on Z_1

equals the vector product of α with the vector of i^{th} partial regression coefficients in the regressions of all of the variables in Z on the subset Z_1.

Now consider the projection of x(t) on the sampled x-process, X(n),

$$\sum_{j=-\infty}^{\infty} \gamma_j^{\,t} X(j) \quad ,$$

where there is one such projection, and hence one sequence $\gamma_j^{\,t}$, for each real t. The regression coefficients $\gamma_j^{\,t}$ are uniquely determined by the orthogonality requirement

$$E[(x(t) - \sum_{j=-\infty}^{\infty} \gamma_j^{\,t} X(j))\, X(n)] = 0$$

or

$$R_x(n-t) = \sum_{j=-\infty}^{\infty} \gamma_j^{\,t} R_X(n-j) \quad .$$

Convoluting both sides of the above equation with R_X^{-1} gives

$$\gamma_n^{\,t} = \sum_{j=-\infty}^{\infty} R_X^{-1}(j)\, R_x(n-t-j)$$

(12)

$$\gamma_n^{\,t} = R_X^{-1} * R_x(n-t) \quad .$$

In equation (12), $\gamma_n^{\,t}$ gives the regression coefficient on X(n) in the projection of x(t) on the X(r) sequence.

Theil's specification formula (11) leads us to expect that

(13) $$B(n) = \int_{-\infty}^{\infty} b(t)\, \gamma_n^{\,t}\, dt.$$

Substituting (12) into the above equation convinces us that the above equation is equivalent with Sims's formula,[6/]

$$B(n) = \int_{-\infty}^{\infty} b(t)\ (R_X^{-1} * R_x(n-t))dt$$

$$= b * R_X^{-1} * R_x(n).$$

Geweke's Formula

The preceding approach can be used to derive Geweke's generalization of Sims's formula where in (1) b(s) is now interpreted as a 1 x k vector at each s, x(t) is a (kxl) vector stochastic process, X(n) is the (kxl) vector discrete time stochastic process corresponding to x(t) sampled at the integers, and the orthogonality condition (2) is modified to be

$$\underset{1x1}{E[u(s)}\ \underset{1xk}{x'(t)]} = \underset{1xk}{0} \qquad \text{for all real t, s.}$$

Now B(s) is 1 x k at each integer s.

The orthogonality condition implies

$$E[(Y(n) - \sum_{s=-\infty}^{\infty} B(s)\ X(n-s))\ X'(r)] = 0 \qquad \text{n, s, r integers}$$

which implies the normal equations

$$(14) \qquad R_{YX}(n-r) = \sum_{s=-\infty}^{\infty} B(s)\ R_X(n-r-s)$$

where $\underset{1xk}{R_{YX}(\tau)} = E[\underset{1x1}{Y(t)}\ \underset{1xk}{X'(t-\tau)}]$

$$\underset{kxk}{R_X(\tau)} = E(\underset{kx1}{X(t)}\ \underset{1xk}{X'(t-\tau)}).$$

Let $R_X^{-1}(j)$ be the (kxk) inverse under convolution of the matrix R_X, where $R_X^{-1}(j)$ is the sequence defined by

$$\sum_{j=-\infty}^{\infty} R_X^{-1}(j)\, R_X(s-j) = \begin{cases} I_{kxk} & s = 0 \\ 0 & s \neq 0 \end{cases} .$$

Convoluting both sides of (14) with R_X^{-1} then gives equation (10), only where all quantities are now interpreted as the matrices defined above,

$$B(\tau) = b * R_X^{-1} * R_x(\tau).$$

This is Geweke's formula.

FOOTNOTES

1/The results do not actually require that b(t) be absolutely integrable. They will remain true if b(t) is viewed as a generalized function, for example, a train of delta functions or derivatives of delta functions. See Sims [ ]. For an introductory discussion of the properties of delta functions and other generalized functions, see Papoulis [ ].

2/That condition (2) uniquely determines the projection is proved, for example, by Ash [p. 121].

3/Again, see Ash [p. 121].

4/Readers familiar with lag operators may find the following helpful. The covariance generating function or z-transform of R_X is

$$\rho_X(z) = \sum_{\tau = -\infty}^{\infty} R_X(\tau) z^{\tau}$$

so that the coefficient on z^{τ} is the covariance at lag τ. The z-transform of the inverse under convolution of R_X, $\rho_X^{-1}(z)$ must satisfy

$$\rho_X(z)\rho_X^{-1}(z) = 1 .$$

Suppose, for example, that x_t follows the moving average process $x_t = B(L)\varepsilon_t$, ε_t white noise with variance σ_ε^2 and $B(L) = (1-b_1 L-b_2 L^2 - \ldots - b_p L^{p})$, where L is the lag operator, $L^n x_t = x_{t-n}$. It is easy to show that $\rho_X(z)$ is given by

$$\rho_X(z) = \sigma_\varepsilon^2 B(z) B(z^{-1}) .$$

(e.g., see Nerlove [ ]). Then the inverse under convolution of R_x has z-transform

$$\rho_X^{-1}(z) = \frac{1}{\sigma_\varepsilon^2} \frac{1}{B(z) B(z^{-1})} .$$

For example, suppose $B(L) = (1-b_1 L)^{-1}$, so that x is first-order Markov. Then

$$\rho_X(z) = \sigma_\varepsilon^2 \frac{1}{(1-b_1 z)} \frac{1}{(1-b_1 z^{-1})} ,$$

$$\rho_X^{-1} = \frac{1}{\sigma_\varepsilon^2} (1-b_1 z) (1-b_1 z^{-1})$$

$$= \frac{1}{\sigma_\varepsilon^2} (-b_1 z^{-1} + (1+b_1^2) - b_1 z) .$$

The value of $R_X^{-1}(n)$ is the coefficient on z^n in the above expression.

5/ Although $R_X^{-1} * R_x(\tau)$ is well defined in the preceding equation of the text, naming it the convolution of a sequence with a function is a slight abuse. More precisely, $R_X^{-1} * R_x(\tau)$ is the convolution of the generalized function, say R_{Xg}^{-1}, corresponding to the sequence R_X^{-1} with R_x. That is, define

$$R_{Xg}^{-1}(t) = \sum_{n=-\infty}^{\infty} R_X^{-1}(n)\ \delta(t-n) ,$$

so that $R_{Xg}^{-1}(t)$ is the generalized function with "mass" $R_X^{-1}(n)$ at integer n, and value zero elsewhere. Then we have

$$\begin{aligned} R_{Xg}^{-1} * R_x(t) = R_x * R_X^{-1}(t) &= \int_{-\infty}^{\infty} R_x(t-\tau) \sum_{n=-\infty}^{\infty} R_X^{-1}(n)\ \delta(t-n)d\tau \\ &= \sum_{n=-\infty}^{\infty} R_X^{-1}(n) \int_{-\infty}^{\infty} R_x(t-\tau)\ \delta(t-n)d\tau \\ &= \sum_{n=-\infty}^{\infty} R_X^{-1}(n)\ R_x(t-n) \end{aligned}$$

which agrees with the definition in the text. In line with the pedagogical purpose of this note, I have kept generalized functions out of the text.

6/ Notice that for t an integer we must have $\gamma_n^t = 1$ for n = t, $\gamma_n^t = 0$ for all $n \neq 0$. This follows because the projection of x(t) on the sequence X(s) is simply X(t) = X(n) for t = n. Since $\gamma_0^t = R_X^{-1} * R_x(-t)$, this shows that the weighting function $R_X^{-1} * R_x(t)$ must be unity at t = 0 and zero at all other integers.

References

Ash, R. A., Real Analysis and Probability, Academic Press, 1972.

Geweke, John, Manuscript, 1974, "Temporal Aggregation in the Multivariate Regression Model," Chapter 6 of "Employment Turnover and Wage Dynamics in U.S. Manufacturing," University of Minnesota Ph. D. Thesis, 1975.

Nerlove, Marc, "Distributed Lags and Unobserved Components in Economic Time Series," in W. Fellner, et al., Ten Economic Studies in the Tradition of Irving Fisher, John Wiley, 1967.

Papoulis, A., The Fourier Integral and Its Applications, McGraw-Hill, 1962.

Sims, C. A., "Approximate Specification in Distributed Lag Models," manuscript, 1971.

Sims, Christopher A., "Discrete Approximations to Continuous Time Distributed Lags in Econometrics," Econometrica, May 1971.

Theil, Henri, Principles of Econometrics, 1971, John Wiley, pp. 548-555.