UNIVARIATE RESIDUAL CROSS CORRELATION ANALYSIS: AN APPLICATION TO BEEF PRICES

by

Steve Miller*

---

*Assistant Professor, Clemson University. Paper presented at American Agricultural Economics Association annual meeting at Washington State University, July 29 - August 1, 1979.

UNIVARIATE RESIDUAL CROSS CORRELATION ANALYSIS: AN APPLICATION TO BEEF PRICES

Introduction

In their empirical work, agricultural economists frequently face questions such as: 1) Does some time-ordered variable X "lead" another time-ordered variable Y? 2) Does feedback exist between X and Y? 3) If Y "lags" X, what is the nature of the lag relationship? The purpose of this study is to discuss and illustrate a relatively new statistical methodology which appears to hold considerable promise as a useful tool for addressing the above questions. While the methodology, univariate residual cross correlation analysis, has been previously employed in assessing macroeconomic lead-lag relationships (e.g.; Pierce), it has yet to be reported in the agricultural economics literature. In subsequent sections this methodology is discussed, and an application is made to beef price changes at the retail, wholesale, and farm levels.

The application is topical owing to recent public concern about pricing efficiency in the beef marketing system. This concern is evidenced in recently proposed federal legislation (The Fair Meat Trading Act of 1979) which would require the Secretary of Agriculture to investigate alternative price discovery mechanisms for beef, including computerized meat trading and a ban on formula pricing (Ward). Since a possible criterion for evaluating alternative price discovery mechanisms might be the speed by which price changes are reflected through the marketing system, analysis of the lead-lag relationships between beef price changes under the current beef marketing system would provide useful "benchmark" information for evaluation of alternative price discovery mechanisms.

Univariate Residual Cross Correlation Analysis

Crucial to the application of univariate residual cross correlation analysis is a notion put forth by Granger. That is, a time-ordered variable X may be said to "lead" or "cause" a second time-ordered variable Y if Y may be better predicted using the past history of X than without, with all other information (including the past history of Y) being used in either case. This criterion has been adapted for use in assessing lead-lag relationships in the time domain by Haugh, and Haugh and Box.

Let Y be generated by a stochastic process, represented in autoregressive and moving average forms as

$$F(B)Y_t = v_t, \tag{1}$$

and

$$Y_t = S(B)v_t, \tag{2}$$

respectively; where Y_t = the value of Y at time t; F(B), S(B) = infinite polynomials in the lag operator B; and v_t = a white noise term at time t with zero mean and constant variance, σ_v^2. 1/ It is assumed that the process is both stationary [S(B) converges for $|B| \leq 1$] and invertible [F(B) converges for $|B| \leq 1$]. 2/

Similarly, let X be generated by a second stochastic process, represented in autoregressive and moving average forms as

$$G(B)X_t = u_t, \tag{3}$$

and

$$X_t = T(B)u_t, \tag{4}$$

respectively; where X_t = the value of X at time t; G(B), T(B) = infinite polynomials in B; and u_t = a white noise term at time t with zero mean and constant variance, σ_u^2. It is assumed that this process is also both stationary and invertible.

As discussed by Haugh and Box, the cross correlation between the u's and v's, defined at lag k as

$$\rho_{uv}(k) = \frac{E\left[u_{t-k}, v_t\right]}{\left[E\left[u_t^2\right] E\left[v_t^2\right]\right]^{1/2}} \quad (5)$$

may be used to assess linear causal or lead-lag relationships between X and Y. Some linear lead-lag relationships of interest, as implied by various patterns in the cross correlations, are shown in Table 1. The reason for examining the cross correlations between the white noise series, rather than the original series, is that any autocorrelation in the original series leads to overestimation of the significance of the cross correlations (Haugh and Box, pp. 122-3).

The u's and v's of Eqs. (1) - (4) are not observable. However, estimates of the u's and v's, denoted as the $\hat{u}$'s and $\hat{v}$'s respectively, may be obtained via application of the univariate time series modeling techniques popularized by Box and Jenkins. These techniques are based on the notion that most time series may be adequately represented by a finite number of autoregressive and/or moving average terms. Statistical tests of the significance of the calculated cross correlations between the $\hat{u}$'s and $\hat{v}$'s, denoted as the $r_{\hat{u}\hat{v}}(k)$'s, may be used to infer the lead-lag relationships between X and Y. If X and Y are independent, the $r_{\hat{u}\hat{v}}(k)$'s are asymptotically, independently and normally distributed with zero mean and variance n^{-1}, where n is the sample size.

As discussed by Pierce, the hypothesis that X and Y are independent may be rejected at significance level α if

Table I. Linear Lead-lag Relationships Implied by Various Cross Correlations

	Cross Correlations at Lag K	Implied Linear Causal Relationships
I.	$\rho_{uv}(k) \neq 0$ for some $k > 0$	X causes Y
II.	$\rho_{uv}(k) \neq 0$ for some $k \leq 0$	Y causes X
III.	$\rho_{uv}(0) \neq 0$	Instantaneous causality between X and Y
IV.	$\rho_{uv}(k) \neq 0$ for some $k > 0$ and for some $k < 0$	Feedback between X and Y
V.	$\rho_{uv}(k) = 0$ for all $k < 0$	Y does not cause X
VI.	$\rho_{uv}(k) = 0$ for all $k > 0$	X does not cause Y
VII.	$\rho_{uv}(k) \neq 0$ for some $k > 0$ and $\rho_{uv}(k) = 0$ for all $k < 0$	Unidirectional causality from X to Y
VIII.	$\rho_{uv}(k) \neq 0$ for some $k < 0$ and $\rho_{uv}(k) = 0$ for all $k > 0$	Unidirectional causality from Y to X
IX.	$\rho_{uv}(k) = 0$ for all $k \neq 0$ and $\rho_{uv}(0) \neq 0$	X and Y are related instantaneously but in no other way
X.	$\rho_{uv}(k) = 0$ for all k	X and Y are independent

Source: (Pierce, p. 15).

$$Q_{2m+1} = n \sum_{k=-m}^{m} \left| r_{\hat{u}\hat{v}}(k) \right|^2 > \chi^2_{\alpha,\ 2m+1} \tag{6}$$

where $\chi^2_{\alpha,\ 2m+1}$ is the upper α percentage point of the chi-square distribution with d.f. = 2m+1; and m is chosen so as to include all $\rho_{uv}(k)$'s expected to differ from zero. The contention that X leads Y is supported at significance level α if

$$Q_m = n \sum_{k=1}^{m} \left| r_{\hat{u}\hat{v}}(k) \right|^2 > \chi^2_{\alpha,\ m} \tag{7}$$

Similarly, Y leads X may be asserted at α if

$$Q_{\bar{m}} = n \sum_{k=-1}^{-m} \left| r_{\hat{u}\hat{v}}(k) \right|^2 > \chi^2_{\alpha,\ m} \tag{8}$$

The significance of an individual $r_{\hat{u}\hat{v}}(k)$ may be determined by comparison to its standard error, $n^{-\frac{1}{2}}$. The convention is to judge an $r_{\hat{u}\hat{v}}(k)$ significant if it is at least twice as large as its standard error. Once the lead-lag relationships are ascertained, the white noise residuals may be used in the construction of dynamic regression models (Haugh and Box).

What are the advantages of univariate residual cross correlation in the assessment of lead-lag relationships between economic time series? Griliches (p. 42) has noted that there is little or no theoretical justification for the specific lag structures often imposed _a priori_ on data. With the present methodology, the analyst need not impose a specific lag structure. Rather, the nature of the lag relationship is an empirical issue; with the time series data themselves being used to assess the relationship. Next, as mentioned above, this approach allows the analyst to avoid testing problems encountered when autocorrelated time series are cross correlated. Similarly, this approach may be used to circumvent the "spurious regression" problem (Granger and Newbold) which would likely be encountered in regressions of Y on leading and/or lagged values of X in order to determine their lead-lag relationships. The "spurious regression" problem; i.e., the overestimation of the significance of both the whole regression and individual regression coefficients, may be encountered when autocorrelated time series are regressed on each other.

With respect to frequency domain techniques, harmonic analysis is sometimes used to assess lead-lag relationships (Franzmann and Walker). However, it is not possible, in a mathematical sense, to distinguish between a lead and a lag by that technique (Barksdale, et al., p. 311). The present methodology does not suffer that problem. While cross-spectral analysis allows assessment of lead-lag relationships, univariate residual cross correlation analysis should be easier to understand and interpret for those not used to frequency domain representations of time series. Also, the present methodology is less of a computational burden than is cross-spectral analysis.

An Application

Univariate residual cross correlation analysis was applied to weekly beef price changes at the retail, wholesale, and farm levels. Lead-lag relationships between price changes at various levels in the beef marketing system have been of recurring interest, owing to previous concern about pricing efficiency in that system. An early study by the National Commission on Food Marketing (pp. 91-94) examined, via regression techniques, weekly price changes. It was concluded that farm price changes led wholesale changes by up to two weeks, and that wholesale price changes led retail price changes by up to eight weeks. The estimated wholesale-retail relationship may, however, have been tainted by a serial correlation problem as indicated by the inconclusive results of a Durbin-Watson test. Franzmann and Walker conducted a harmonic analysis of monthly wholesale and farm level prices, and concluded that wholesale prices led farm prices by three months. Barksdale, et al. used cross-spectral techniques in an analysis of monthly prices. Their results indicated that wholesale and farm prices changed instantaneously, but that retail prices followed wholesale prices by three weeks. Most recently, King used polynomial distributed lags in an analysis of weekly price changes. He concluded that farm and wholesale prices were instantaneously related, and that wholesale prices led retail prices by up to five weeks. King acknowledged that his models suffered from certain estimation problems, particularly, serially correlated residuals.

The data for the present analysis were comprised of the first differences of weekly retail, wholesale, and net farm beef values for January, 1974 through June, 1978, with n = 234.[3/] Let R_t, W_t, F_t = the first differences of retail, wholesale, and net farm beef values, respectively, between weeks t and t-1. Table 2 shows the estimated autocorrelations of the R's, W's and F's for up to 10 lags. If each series was random, the standard error of individual autocorrelations could be approximated by $n^{-\frac{1}{2}}$, which in the present case is $234^{-\frac{1}{2}} = .07$. Note that each series in Table 2 has estimated autocorrelations exceeding the value .07 by a factor of two or more. It may be concluded that autocorrelation is present in each series. Recall from the previous section that this autocorrelation may lead to overestimation of the significance of cross correlations between R's and W's, and/or the W's and F's. Thus, the univariate residual cross correlation approach is appropriate.

The iterative model building process described by Box and Jenkins resulted in the following univariate time series models:

Retail: $\hat{a}_t = R_t - .11386\hat{a}_{t-1} - .16216\hat{a}_{t-3} = .23175\hat{a}_{t-4}$, RSE = 1.86 (9)

wholesale: $\hat{b}_t = W_t - .15320\hat{b}_{t-3}$, RSE = 3.11 (10)

and farm: $\hat{c}_t = F_t - .26258\hat{c}_{t-1} - .29304\hat{c}_{t-2}$, RSE = 2.54 (11)

where the $\hat{a}$'s, $\hat{b}$'s and $\hat{c}$'s = white noise residuals; and RSE = the residual standard error.[4/] Calculated cross correlations between the $\hat{a}$'s and $\hat{b}$'s and the $\hat{b}$'s and $\hat{c}$'s for $-10 \leq k \leq 10$ are reported in Tables 3 and 4.

From the Q-statistics (with m chosen to be 10), and comparison of individual cross correlations to their standard errors, the following points appear noteworthy.

1. In Table 3, $Q_{\overline{10}}$ exceeds the critical value of $\chi^2_{5,10} = 18.3$, but Q_{10} does not. The implication is that wholesale value changes preceed retail value changes, and there is no feedback from the retail level. The largest cross correlations occur at the first two negative lags of $\hat{a}_t$, implying a more rapid adjustment of retail prices than that found by King and the National Commission on Food Marketing. However, the serial correlation problems suffered in those studies may have resulted in the "spurious regression" phenomenon discussed by Granger and Newbold.

2. From Table 4, $Q_{\overline{10}}$ exceeds the critical value of 18.3, while Q_{10} does not. Large cross correlations occur at $r_{\hat{b}\hat{c}}(0)$ and $r_{\hat{b}\hat{c}}(-1)$. These results imply that farm value changes lead wholesale value changes, lending some support to the findings of the National Commission on Food Marketing, but contradicting Barksdale, _et al._, and King, who concluded that wholesale and farm level changes were instantaneous. However, farm and wholesale value changes had their strongest association at a zero lag in the present analysis. These results sharply contradict the conclusion

Table 2. Estimated Autocorrelations of Weekly Price Changes at the Retail, Wholesale, and Farm Levels.

Lags									
1	2	3	4	5	6	7	8	9	10
				Retail Level					
.20	.15	.17	.22	.11	.02	-.04	.00	.05	-.03
				Wholesale Level					
.08	-.02	.15	-.02	-.05	.00	.02	-.05	-.05	.02
				Farm Level					
.28	.21	-.04	-.06	-.01	-.04	-.03	-.05	-.03	.02

Table 3. Estimated Cross Correlations Between White Noise Residuals of Weekly Retail and Wholesale Value Changes[a]

Lags									
1	2	3	4	5	6	7	8	9	10
				Positive Lags of $\hat{a}_t$ $(k > 0)$					
-.10	.11	.02	-.01	-.06	-.04	.01	.01	-.02	-.02
				Negative Lags of $\hat{a}_t$ $(k < 0)$					
.46*	.43*	.12	.01	-.08	.05	.05	.05	.03	.03

Note: $r_{\hat{a}\hat{b}}(0) = .10$, $234^{-\frac{1}{2}} = .07$, $Q_{21} = 109.4$, $Q_{10} = 6.8$, $Q_{\overline{10}} = 100.3$

Table 4. Estimated Cross Correlations Between White Noise Residuals of Weekly Wholesale and Net Farm Value Changes[a]

Lags									
1	2	3	4	5	6	7	8	9	10
				Positive Lags of $\hat{b}_t$ $(k > 0)$					
.04	-.07	.02	-.08	.05	.03	.00	-.08	-.02	.02
				Negative Lags of $\hat{b}_t$ $(k < 0)$					
.22*	.12	-.08	-.02	.01	-.01	.04	-.03	-.07	.09

Note: $r_{\hat{b}\hat{c}}(0) = .76^*$, $Q_{21} = 161.4$, $Q_{10} = 5.6$, $Q_{\overline{10}} = 20.0$

[a] In Tables 3 and 4, a cross correlation marked by * is at least three times greater than its standard error.

of Franzmann and Walker that wholesale prices lead farm prices by three months.

Taken as a whole, the present analysis indicates farm level changes are apparently reflected in wholesale value changes within a week, and wholesale changes are reflected in retail value changes within three weeks. Considering the two weeks or so required to physically transform a live beef animal into beef cuts at the retail level (Faris and Couvillion), these results imply that the current price discovery mechanisms in the beef marketing system provide fairly rapid price adjustments between the farm, wholesale, and retail levels.

While the above analysis has indicated the lead-lag relationships between price changes in the beef marketing system, the analysis has not indicated what proportion of farm level changes are transmitted to the wholesale level, or the proportion of wholesale level changes transmitted to the retail level. An approach to that problem would be to use the lead-lag relationships identified above in the construction of dynamic regression models explaining wholesale level changes with farm level changes, and retail level changes with wholesale level changes. The author is presently constructing such models.

Summary

This study has provided a discussion of univariate residual cross correlation analysis, a statistical tool for assessing lead-lag relationships between economic time series. This methodology offers several advantages. The analyst need not *a priori* impose a specific lag structure on his data; rather, the data are allowed to suggest the nature of the lead-lag relationships. Univariate residual cross correlation analysis is also both easier to understand and to apply than is cross-spectral analysis.

For the purpose of illustration, this methodology was applied in an analysis of lead-lag relationships between beef value changes at the retail, wholesale, and farm levels. The results of that analysis indicate that farm prices lead wholesale prices by about one week, and in turn, wholesale prices lead retail prices by about three weeks. These results were interpreted as being indicative of a relatively rapid adjustment of beef prices through the beef marketing system.

Footnotes

1/ The lag operator is defined such the $B^j v_t = v_{t-j}$, for example.

2/ Stationarity implies that the probabilistic properties of the process are not affected by a shift in time origin. A non-stationary process can often be made stationary by a judicious data transformation. Invertibility assures that v_t may be reproduced or recovered from present and past values of Y.

3/ The data are those reported by the Meat Animals Program Area, USDA, ESCS, CED.

4/ The residuals were deemed white noise based on the absence of any patterns among their AC's and PAC's. This conclusion was supported by the results of the "portmanteau" test, with $\alpha = .05$, suggested by Box and Jenkins. Ninety-five percent confidence limits for the parameters associated with the above variables were as follows: $\hat{a}_{t-1}$, -.24 to .01; $\hat{a}_{t-3}$, -.29 to -.03; $\hat{a}_{t-4}$, -.36 to -.10; $\hat{b}_{t-3}$, -.28 to -.02; $\hat{c}_{t-1}$, -.39 to -.13; $\hat{c}_{t-2}$, -.42 to -.17.

References

Barksdale, H. C., J. E. Hilliard, and M. C. Ahlund. "A Cross-Spectral Analysis of Beef Prices," American Journal of Agricultural Economics. 57(1975): 309-315.

Box, G. E. P. and G. M. Jenkins. Time Series Analysis: Forecasting and Control. (Revised Edition). San Francisco: Holden-Day, 1976.

Faris, D. E. and W. C. Couvillion. Vertical Coordination of Beef in the South - Nature of Different Systems. Southern Cooperative Series Bulletin 192, 1975.

Franzmann, J. R. and R. L. Walker. "Trend Models of Feeder Slaughter, and Wholesale Beef Cattle Prices," American Journal of Agricultural Economics. 54(1972): 507-512.

Granger, C. W. J. "Investigating Causal Relations by Econometric Models and Cross Spectral Methods," Econometrica. 37(1969): 424-438.

Granger, C. W. J. and P. Newbold. "Spurious Regressions in Econometrics," Journal of Econometrics. 2(1974): 111-120.

Griliches, Z. "Distributed Lags: A Survey," Econometrica. 37(1967): 16-49.

Haugh, L. D. "Checking the Independence of Two Covariance-Stationary Time Series: A Univariate Residual Cross Correlation Approach," Journal of the American Statistical Association. 71(1976): 378-385.

Haugh, L. D. and G. E. P. Box. "Identification of Dynamic Regression (Distributed Lag) Models Connecting Two Time Series," Journal of the American Statistical Association. 72(1977): 121-130.

King, R. A. "Transmission of Week-to-Week Change in Choice Beef Prices Between Farm, Carcass and Retail Levels," Contributed Paper, AAEA meeting at Penn. State University, 1976.

National Commission on Food Marketing. Organization and Competition in the Livestock and Meat Industry. Technical Study No. 1, 1966.

Pierce, D. A. "Relationships - and the Lack Thereof - Between Economic Time Series, with Special Reference to Money and Interest Rates," Journal of the American Statistical Association. 71(1977): 11-22.

Ward, C. E. "Proposed Meat Marketing and Pricing Legislation," Agricultural Policy and Economic Issues. Oklahoma Cooperative Extension Service, April, 1979.