Modeling Advertising Carryover in Fluid Milk: Comparison of Alternative Lag Specifications

Meenakshi Venkateswaran, Henry W. Kinnucan, and Hui-Shung Chang

The performance of restricted estimators such as Almon and Shiller in modeling advertising carryover is tested and compared to the unrestricted OLS estimator, using 1971–1988 monthly New York City fluid milk market data. Results indicate that in the absence of autocorrelation and multicollinearity among the lagged advertising variables, the unrestricted OLS estimator is still the preferred estimator, based on Mean Square Error and Root Mean Square Percent Error criteria. In this case, the Almon and Shiller estimators perform equally well, although next only to the OLS estimator. In the presence of autocorrelation or multicollinearity however, the restricted estimators may outperform the OLS estimator, in a MSE sense, with the flexible Shiller estimator (which subsumes the Almon) being more desirable.

Advertising carryover, a phenomenon perhaps best described by Fred Waugh's aphorism (p. 367) ". . . old advertisements never die—they just fade away," is an issue in current economic research concerning producer-funded advertising campaigns (e.g., see Chang and Kinnucan; Liu and Forker; Ward and Dixon). Reliable estimates of the advertising lag structure are pivotal for accurate assessment of optimal spending levels, producer returns, optimal fund allocations, and other issues. In distributed lag models of advertising response, multicollinearity among the lagged independent variables, if serious, leads to imprecise estimates of the parameters obtained from unrestricted Ordinary Least Squares (OLS). The Almon procedure is a widely used alternative in such situations.

The Almon procedure requires that the lag weights lie exactly on a polynomial of known degree over a known interval (Almon). This procedure has been criticized as being too restrictive because the shapes of the estimated lag structures are dictated by the stringency of specifications (Fomby). In an extensive Monte Carlo study, Cargill and Meyer found that the Almon restrictions produced severely distorted shapes of the lag distribution. Their results showed that Almon estimates for a second degree polynomial, whether the end points are constrained or not, yielded biased estimates, which in many cases exceeded 50 percent of the true values of the respective coefficients. While increasing the polynomial to a fourth degree reduced the size of the biases, they were still large compared to unrestricted OLS estimates. Misspecification of the lag length and the presence of serial correlation tended to increase the biases further.

A technique, which imposes fewer restrictions and includes the unrestricted OLS and Almon procedures as limiting cases, has been developed by Shiller. The Shiller procedure is more flexible in that the lag structure follows a polynomial of a given degree not exactly but stochastically. In other words, restrictions on the lag structure are made stochastic by requiring the means of the lag coefficients to lie on a polynomial of certain degree, while allowing the coefficients to vary smoothly.

Shiller evaluated lag distributions of known shapes using OLS, Almon and Shiller procedures. Results showed OLS producing a jagged representation of the true shape. The Almon estimates, in general, did a poor job of representing the tails of the true distribution. The Shiller estimates, by con-

Post-doctoral research fellow, associate professor, and former visiting assistant professor, respectively, Department of Agricultural Economics and Rural Sociology, Auburn University.

Appreciation is expressed to Patricia Duffy, Robert Nelson, and Greg Traxler for reviewing an earlier draft of this manuscript. The usual disclaimers apply.

Scientific Journal Paper No. 1-91939A of the Alabama Agricultural Experiment Station.

trast, produced a smooth-shaped distribution that closely mimicked the true shape. Fomby applied the Shiller methodology to the data sets used by Almon, and Griliches et al. For the Almon data, when a polynomial of degree two was chosen, both the Almon and Shiller procedures produced estimates with a lower Mean Squared Error (MSE) than the OLS estimates. For the Griliches et al. data with a first degree polynomial restriction on the Almon estimator, the hypothesis of MSE superiority of the Almon over the OLS estimates was rejected.

A number of studies designed to determine the economic effectiveness of generic milk advertising have employed the Almon procedure in estimating the relationship between milk sales and advertising expenditures (Kaiser et al; Thompson and Eiler; Ward and Dixon). The purpose of this paper is to determine the performance of the Almon and the less restrictive Shiller procedure in modeling advertising carryover. Performance is evaluated using the Mean Square Error (MSE) and *ex post* Root Mean Square Percent Error (RMS%E) criteria.

The paper begins with a brief review of the Almon distributed lag model and the procedure used to estimate the Almon coefficients. The Shiller procedure and its estimation procedures are discussed next. In this section we introduce the Theil-Goldberger (TG) mixed estimation procedure traditionally used to obtain Shiller estimates and then discuss a new technique, the Prior Integrated Mixed Estimation (PIME) procedure introduced by Mittelhammer and Conway, and its superiority over the TG estimator. The MSE and RMS%E criteria used to evaluate the different specifications are then explained. We then present the empirical model and data and discuss our results. Some concluding comments about using the alternative lag specifications are presented in the final section of the paper.

The Almon Distributed Lag Model

The idea underlying the distributed lag models of finite lengths is that the researcher generally has some *a priori* notions about the likely appearance of the lag coefficients and that these notions should be explicitly incorporated into the estimation framework, in the form of restrictions, to increase the precision of the estimates. In the Almon polynomial distributed lag model the belief is that the lag coefficient should lie exactly on a polynomial of a chosen degree. Specifically, suppose the model is

$$(1)\quad y_t = \sum_{i=0}^{n} \beta_i x_{t-i} + \xi_t, \qquad \xi_t \sim N(0, \sigma^2)$$

where y_t and x_t are elements of a scalar time series of the dependent and the independent variables, respectively, a time t (t = 1, . . . T). One way of imposing the restriction that the lag coefficients lie exactly on a polynomial of degree p is by imposing (n − p) restrictions of the form

$$(2)\quad \Delta^{p+1} \beta_i = 0, \qquad \text{for } i = p + 1, \ldots n$$

where Δ is the difference operator (e.g., $\Delta \beta_i = \beta_i - \beta_{i-1}$), and n is the lag length.

Estimation of the Almon Coefficients

The Almon estimates are obtained by solving a restricted least squares problem where the constraints are given by (2) above. A simple way to compute the restricted least squares estimate for β is to replace the β vector by $H_p\alpha_p$ where $\alpha_p = (\alpha_0, \alpha_1, \ldots \alpha_p)'$ and

$$(3)\qquad H_p = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 1 & 1 & 1 & \ldots & 1 \\ 1 & 2 & 2^2 & \ldots & 2^p \\ . & . & . & \ldots & . \\ . & . & . & \ldots & . \\ 1 & n & n^2 & \ldots & n^p \end{bmatrix}$$

Hence we have a linear model

$$(4)\qquad Y = X H_p\alpha_p + \xi = Z \alpha_p + \xi,$$

where $Z = X H_p$. Using this model, the least squares estimator for α_p is

$$(5)\qquad \hat{\alpha}_p = (Z'Z)^{-1} Z'Y$$

and the restricted least squares estimator of β is

$$(6)\qquad \hat{\beta} = H_p \hat{\alpha}_p = H_p(Z'Z)^{-1} Z'Y$$

The Shiller Distributed Lag Model

The basic assumption of the Shiller lag model is that the lag coefficients trace a smooth curve around the chosen polynomial. The Shiller approach provides a flexible means of incorporating the smoothness assumption in estimations of distributed lag models of finite lengths. For the distributed lag model described by (1), the smoothness restriction is imposed by requiring

$$(7) \qquad \Delta^{p+1} \beta_i = \omega_i, \qquad \omega_i \sim N(0, \phi^2).$$

p in this case is called the degree of the smoothness prior. For example, a second degree smoothness prior requires the third differences in the β_i to be small or approximately zero for all i. The prior information in this case is specified in probabilistic terms by making the smoothness restrictions stochastic. Imposing "smoothness" priors is therefore equivalent to imposing stochastic, rather than exact restrictions. Thus, ϕ^2 reflects the degree of confidence in the researchers' prior beliefs.

Estimation of Shiller Coefficients

Theil-Goldberger Mixed Estimation

The mixed estimation procedure introduced by Theil and Goldberger is a convenient way of incorporating Shiller's smoothness restrictions into the estimation procedure (Fomby; Taylor). The framework consists of two sets of stochastic linear equations involving the k × 1 vector of model parameters, β, denoted by

$$(8) \qquad \begin{bmatrix} Y \\ r \end{bmatrix} = \begin{bmatrix} X \\ R \end{bmatrix} \beta + \begin{bmatrix} \xi \\ \omega \end{bmatrix}$$

where Y and ξ are (n × 1) vectors, r and ω are (j × 1) vectors, X and R are known (n × k) and (j × k) matrices of ranks k and j, respectively. $E[\xi] = [0]$, $\text{cov}[\xi] = \sigma^2 I$, $E[\omega] = \delta$, $\text{cov}[\omega] = \Psi$ is a positive definite matrix, and ξ and ω are independent. A basic assumption of mixed estimation is that r must be non-degenerate stochastic, i.e., have a non-zero variance.

The first set of n equations in (8) is the usual linear model representation of sample information, and the second set of j equations refers to the prior introspective information. The vector ω represents the deviations of the outcomes of r from the vector $R\beta$ (Mittelhammer and Conway). Using the Generalized Least Squares (GLS) procedure, the TG estimator is defined as

$$(9) \qquad \hat{\beta}_{TG} = (\sigma^{-2} X'X + R' \Psi^{-1} R)^{-1} (\sigma^{-2} X'Y + R' \Psi^{-1} r).$$

The TG mixed estimation procedure, however, has recently been criticized for lacking theoretical basis (Swamy and Mehta; Judge et al.) and violating some basic assumptions of mixed estimation (Mittelhammer and Conway). Mittelhammer and Conway indicate the source of the problem to be the genesis and interpretation of the prior introspective estimator r. In the TG framework, r is of the subjectivist type where the researcher formulates an *a priori* point estimate, as well as its sampling variance (Theil and Goldberger, p. 73), based on best guesses of the corresponding elements of $R\beta$ (Theil). Swamy and Mehta point out that r in this case must be degenerate stochastic, or in other words, have zero variance, since a researcher's best guess is in general a fixed constant.

The degenerate stochasticity of r in the TG estimator contradicts the basic assumption of mixed estimation, which requires r to be non-degenerate stochastic. In the Shiller procedure, r is always set equal to zero and therefore is degenerate stochastic. Hence, it is inappropriate to use the TG mixed estimator to estimate Shiller lags. Similarly, Fomby's procedure to test the null hypotheses of MSE superiority of mixed estimators is not applicable since the test procedure is based on the assumption of non-degenerate stochastic distribution of r.

Prior Integrated Mixed Estimation

Mittelhammer and Conway provide an alternative—the Prior Integrated Mixed Estimator (PIME)—that overcomes the problem encountered in using the TG mixed estimator. This procedure involves expressing the prior information on $R\beta$ in terms of a subjective probability distribution of potential $R\beta$ values with mean vector Ω and covariance matrix Ψ. The subjective distribution assigned to r represents the perceived relative degrees of belief about the correctness of various estimates of the model. This explicit distributional assumption, in addition to the standard assumptions underlying (8), forms the conceptual base for the PIME. The prior integrated mixed estimator of β is given by

$$(10) \qquad \hat{\beta}_{PIME} = (\sigma^{-2} X'X + R' \Psi^{-1} R)^{-1} (\sigma^{-2} X'Y + R' \Psi^{-1}\Omega).$$

It may be noted that $\hat{\beta}_{PIME}$ is the same as $\hat{\beta}_{TG}$ except that r in the stochastic restriction $r = R\beta + \omega$ is replaced by Ω. The PIME has two advantages: (i) It has a smaller mean square error matrix than the TG mixed estimator, and (ii) it is consistent with Theil's notion of incorporating a researcher's "best guesses" and with Swamy and Mehta's argument that the best guess is a constant.

The subjective nature of the probability distribution in the PIME framework subsumes the TG mixed estimator as a special case and allows the legitimate use of zero in place of r, as required in the Shiller procedure. Therefore, the estimated values of β according to the Shiller procedure are

$$(11) \quad \hat{\beta}_s = (\sigma^{-2} X'X + R'\Psi^{-1}R)^{-1} (\sigma^{-2}X'Y).$$

If we assume the stochastic restrictions on β_i to be independent with a common variance ϕ^2, then $\Psi = \phi^2 I$. Equation (11) can therefore be rewritten as

$$(12) \quad \hat{\beta}_s = (X'X + k R'R)^{-1} (X'Y).^1$$

In (12), k is a tightness parameter given by σ^2/ϕ^2, where σ and ϕ are known *a priori*. This parameter indicates the relative confidence of the researcher on the sample versus prior information. OLS estimates can be obtained from the Shiller estimator when k tends to zero. Likewise, when k tends to infinity, the Shiller estimator based on a second degree polynomial, results in Almon estimates (Shiller).

The Shiller methodology requires an estimate for k when σ and ϕ are unknown. For $d = 1$, Shiller suggests a rule-of-thumb which involves setting $\phi = 8(s/n^2)$ where s is the sum of the lag coefficients obtained from OLS and n is the lag length. The value of k computed in this manner has a limitation in that it is not invariant to changes in the units of measurement of the variables. Maddala suggests an alternate procedure to compute k. This procedure, developed by Lindley and Smith, involves iteratively estimating k starting with the OLS estimates of the lagged variables and taking the variance of these estimates as a primary estimate of ϕ^2. Shiller's estimates are then obtained using $\hat{k} = \hat{\sigma}^2/\hat{\phi}^2$, where $\hat{\sigma}^2$ is the residual variance from the OLS regression. In the next step, the estimate of $\hat{\phi}^2$ is revised using the residual variance of the Shiller's procedure, to obtain a new value for k and the procedure is repeated until a satisfactory level of convergence is attained. The Maddala estimator of k in this case is similar to the Hoerl-Kennard estimator of the biasing parameter in ridge regression.

The Evaluation Procedures

The Mean Square Error (MSE) Criterion

Imposing restrictions on the parameters increase the precision (i.e., reduces the variance) of parameter estimates but, unless correct, the restrictions produce biased estimates. The goodness of the parameters are assessed using the MSE criterion. The MSE for a given parameter β_i, as

$$(13) \quad \text{MSE}(\hat{\beta}_i) = \text{Variance}(\hat{\beta}_i) + (\text{Bias}(\hat{\beta}_i))^2$$

where $\hat{\beta}_i$ is the estimated value of the true parameter β_i. According to this criterion, an estimator producing the smallest MSE would be preferred.

In practice, we do not know the value of the true parameter β_i. Hence, we cannot estimate the MSE. However, we can test the hypothesis of MSE superiority of one estimator over the other. Applying the MSE criterion to the Shiller estimator we can form the hypotheses

$$(14a) \quad H_N: E[(\hat{\beta}_s - \beta)'(\hat{\beta}_s - \beta)] \leq E[(\hat{\beta}_o - \beta)'(\hat{\beta}_o - \beta)]$$

$$(14b) \quad H_A: H_N \text{ not true}$$

where $\hat{\beta}_s$ and $\hat{\beta}_o$ are the Shiller and OLS estimators respectively. Rejection of H_N implies the Shiller estimator is not superior to the OLS estimator in a MSE sense.

The Strong Mean Square Error (SMSE) superiority[2] of Shiller estimates over the OLS estimates, for instance, can be tested using the *F*-statistic introduced by Mittelhammer and Conway. The test statistic in this case is computed as

$$(15) \quad f = j^{-1}(R\hat{\beta}_o - \Omega)'(s^2R(X'X)^{-1}R')^{-1}(R\hat{\beta}_o - \Omega),$$

where

$$(16) \quad s^2 = (Y - X\hat{\beta}_o)'(Y - X\hat{\beta}_o)/(n - k).$$

Defining $F_\alpha(j,(n-k); 1/2)$ to be the α-level upper tail critical point of a non-central *F*-distribution of j and $(n - k)$ degrees of freedom, and non-centrality parameter 1/2, the specific test is

$$(17) \quad \text{if } f \begin{bmatrix} \leq \\ > \end{bmatrix} F_\alpha(j, (n-k); 1/2), \text{ then } \begin{bmatrix} \text{do not reject} \\ \text{reject} \end{bmatrix} \hat{\beta}_{PIME} \text{ SMSE superiority.}$$

Tabulated values of F_α are provided in Wallace and Toro-Vizcarrondo.

[1] As Maddala points out, the Shiller estimator is also a ridge estimator (Hoerl and Kennard).

[2] An estimator b_1 is Strong Mean Square Superior (SMSE) to b_2 when the mean square error matrix of b_2 exceeds the mean square error matrix of b_1 by a positive semidefinite matrix (Mittelhammer and Conway, p. 862).

The ex post Root Mean Square Percent Error (RMS%E) Criterion

The MSE test discussed above is helpful in testing whether the Almon and Shiller specifications are superior to the OLS estimator. The test, however, does not permit direct comparison of the two restricted estimators—Shiller and Almon. This is a major limitation of the above test. The RMS%E is an alternative criterion that can be used to compare the performance of all three estimators directly.[3] Although this criterion does not involve a structured test, it provides a three way comparison and allows one to observe how well the data fit all three models under consideration. The RMS%E computed as

$$(18)\quad \sqrt{\frac{1}{N - T} \sum_{t=T+1}^{N} \left[\frac{Y^p_t - Y^a_t}{Y^a_t} * 100\right]^2}$$

measures the *ex post* percent deviation of the predicted values of the dependent variable from its actual value.

For evaluation purposes, both in-sample and out-of-sample *ex post* forecasts were obtained. For the in-sample forecast ($t = 1, 2, \ldots, T$), the predicted value of the dependent variable (Y^p_t) was obtained for the period July 1971 through December 1988, using the actual data on explanatory variables for the said period. For the out-of-sample forecast, on the other hand, values of the dependent variable are predicted beyond the estimation period (i.e., $t = T + 1, \ldots, N$) using each of the three estimators and the actual monthly data for the explanatory variables from January 1989 through December 1990. These *ex post* forecasts are then checked against the actual values of the dependent variable (Y^a_t) for the corresponding period. The estimator with the smallest RMS%E is the preferred one.

Empirical Model and Data

Demand for fluid milk is affected by its own price, prices of related beverages, consumers' incomes, current and past advertising efforts, demographics and seasonality. We therefore model fluid milk sales response as a function of the above explanatory variables. The function is specified in double log form to permit diminishing marginal returns to advertising (Simon and Arndt; Venkateswaran and Kinnucan). Algebraically, the reduced-form quantity determination model used is

$$(19)\quad \ln Q_t = \ln \alpha + \lambda \ln PM_t + \epsilon \ln PC_t + \rho \ln PCF_t + \theta \ln I_t + \nu T_t + \sum_{j=1}^{11} \gamma_j Z_{jt} + \sum_{i=0}^{n} \beta_i \ln A_{t-i} + \xi_t$$

where Q_t refers to per capita daily milk sales in quarts in month t ($t = 7, \ldots, 216$, corresponding to July 1971 through December 1988), PM is the real price of milk ($/quart), and PC and PCF are the real price indexes of cola and coffee respectively; I denotes average weekly personal income before taxes, measured in real dollars. T is a time trend variable included to capture the effect of systematic and/or time related changes in demographics and secular growth or decline in milk sales in New York City over time (Kinnucan, 1986). The Z_j's are eleven (0,1) monthly seasonality dummy variables with December being the base month.

The A_{t-i} are current and lagged values of real monthly per capita generic fluid milk advertising expenditures; and n is the lag length. Previous studies (Thompson and Eiler; Liu and Forker; Kinnucan and Forker) indicated a lag length of six months is appropriate to measure the carryover effect of fluid milk advertising in New York City. The value of n in equation (19), therefore, is set equal to six.

ξ is the error term and the other Greek letters refer to the regression coefficients to be estimated. All variables except the seasonal dummies and time trend are expressed in natural logarithms. The income and price variables are deflated using CPI (1981 = 100). The advertising expenditure data are the actual (not budgeted) total expenditures for all media, i.e., T.V., radio, print, and outdoor advertising. A media cost index specific to the New York City media coverage area is used to deflate advertising expenditures.

The data for New York City milk sales and advertising expenditures were obtained from records kept by the New York State Department of Agriculture and Markets. The price, income, and population data for the study area were obtained from government statistics. The media cost index figures were obtained from D'Arcy, Masius, Benton and Bowles (DMB&B)—the advertising agency handling the New York milk account. A data appendix containing specific references for data sources is available upon request from the authors.

[3] The authors are grateful to an anonymous journal reviewer for suggesting this test.

Empirical Results

The OLS, Almon, and Shiller estimates of the New York City fluid milk demand model are presented in Table 1. All three models provide good explanatory power in that the R^2 is 80% or better. The estimated coefficients for all three estimators have the correct signs and are significant in most cases. The own-price coefficient is negative and significant in all three cases indicating a price inelastic demand for fluid milk. Cola and coffee appear to be gross substitutes for fluid milk under all three estimators. The own- and cross-price elasticities are consistent with the elasticities obtained by Kinnucan *et al.* for an earlier period. Income has a significant influence on milk demand as expected. The income elasticity in each of the three estimators is about 0.27, indicating the demand for fluid milk is income inelastic. This finding is consistent with previous milk demand studies (Liu and Forker; Kinnucan, 1986).

Time trend is negative and significant in all three estimators indicating a trend-related decline in milk sales over time. This result is consistent with Kinnucan's (1986) finding that as the proportion of older people in the population increases, milk consumption declines. The estimated coefficients of the monthly seasonality dummies indicate a significant reduction in milk sales from May through September (i.e., summer and fall) and during November compared to December. The reduction in milk consumption during summer and fall can be attributed to the increased consumption of soft drinks, sodas and juices during these months. Similarly, the increased consumption of apple cider, eggnog, and other beverages at Thanksgiving may explain the lower milk consumption in November.

The estimated coefficients of current and lagged advertising expenditures are positive in all the models considered. They are all significant at the five per cent level or better in all but one case in the OLS estimator. The magnitudes of the standard errors of the advertising coefficients are as expected—the lowest in the case of the Almon estimator, followed closely by the Shiller and OLS estimators, in that order. The long-run advertising elasticity, obtained as the sum of the current and lagged coefficients ranges between 0.0133 and 0.0135 for the three estimators. These estimates are smaller than the estimates obtained for New York City based on earlier data (Kinnucan and Forker; Kinnucan, 1986) but are consistent with the findings of Ward and Dixon for other major milk markets in the United States.

The performance of the Shiller estimator is as-

Table 1. Estimated Coefficients of the New York City Fluid Milk Demand Model: OLS, Almon, and Shiller Estimates, Untransformed Data.

	Estimated regression coefficients obtained using		
	Unrestricted estimator	Restricted estimators	
Variables	OLS	Almon	Shiller
Intercept	−2.7086*	−2.7351*	−2.7335*
	(0.4095)@	(0.4145)	(0.4081)
Price of milk	−0.0802*	−0.0787*	−0.0791*
	(0.0384)	(0.0390)	(0.0384)
Price of cola	0.1368*	0.1380*	0.1382*
	(0.0279)	(0.0283)	(0.0279)
Price of coffee	0.0389*	0.0386*	0.0387*
	(0.0061)	(0.0062)	(0.0061)
Income	0.2673*	0.2732	0.2729
	(0.0929)	(0.0941)	(0.0926)
Trend	−0.0003*	−0.0003*	−0.0003*
	(0.0001)	(0.0001)	(0.0001)
Dummy variable for:			
January	0.0036	0.0037	0.0035
	(0.0068)	(0.0068)	(0.0067)
February	−0.0041	−0.0023	−0.0033
	(0.0070)	(0.0068)	(0.0069)
March	0.0080	0.0082	0.0079
	(0.0069)	(0.0068)	(0.0068)
April	−0.0098	−0.0109	−0.0106
	(0.0071)	(0.0071)	(0.0070)
May	−0.0201*	−0.0200*	−0.0203*
	(0.0069)	(0.0069)	(0.0069)
June	−0.0357*	−0.0332*	−0.0339*
	(0.0069)	(0.0069)	(0.0069)
July	−0.0971*	−0.0971*	−0.0974*
	(0.0072)	(0.0071)	(0.0071)
August	−0.0882*	−0.0879*	−0.0884*
	(0.0077)	(0.0077)	(0.0076)
September	−0.0195*	−0.0190*	−0.0193*
	(0.0073)	(0.0073)	(0.0073)
October	0.0011	−0.0015	0.0013
	(0.0069)	(0.0070)	(0.0069)
November	−0.0122*	−0.0122*	−0.0125*
	(0.0066)	(0.0067)	(0.0066)
Advertising expenditure at time			
t	0.0024*	0.0018*	0.0020*
	(0.0006)	(0.0004)	(0.0005)
t-1	0.0010	0.0019*	0.0017*
	(0.0006)	(0.0003)	(0.0004)
t-2	0.0026*	0.0020*	0.0019*
	(0.0006)	(0.0002)	(0.0003)
t-3	0.0014*	0.0020*	0.0022*
	(0.0006)	(0.0003)	(0.0003)
t-4	0.0033*	0.0020*	0.0021*
	(0.0006)	(0.0002)	(0.0003)
t-5	0.0004	0.0019*	0.0017*
	(0.0006)	(0.0003)	(0.0004)
t-6	0.0024*	0.0017*	0.0018*
	(0.0006)	(0.0004)	(0.0005)
Sum of the advertising coefficients	0.0135	0.0133	0.0134
Adjusted R^2	0.8188	0.8129	0.8093
Durbin-Watson Statistic	1.3350	1.4433	1.4282

@: Figures in the parentheses are standard errors.
*: Statistically significant at the 5 percent level.

sessed using Mittelhammer and Conway's F-test, which tests the hypothesis of SMSE superiority of the Shiller over the OLS estimates. The F-statistic so computed (2.56), is less than the critical F(2.97), as shown in table 3. Hence, we fail to reject the null hypothesis of SMSE superiority of the Shiller estimator over the unrestricted OLS estimator for these data. We also compared the performance of the Almon estimator over OLS using the same test. Our null hypothesis in this case was that the Almon estimator is SMSE superior to the OLS estimator. The computed F(2.56) is again less than the critical value (2.97).[4] The null hypothesis of SMSE superiority of the Almon estimator over OLS is not rejected here as well. These findings imply that the Almon and Shiller estimators are superior to the unrestricted OLS estimator, in the MSE sense.

The relative performance of the Shiller and Almon estimators cannot be directly compared using the above F-test since the former involves stochastic linear restrictions whereas the latter involves exact linear restrictions. The problem was overcome by using the RMS%E test, which permits a three-way comparison of the estimators. RMS%E's were computed for both in-sample and out-of-sample *ex post* forecasts for all three estimators and are shown in table 3.

The *ex post* RMS%E for the out-of-sample forecast ranged from 4.86 in the case of OLS to 4.94 in the case of Shiller. The RMS%E of the Almon estimator (4.93) was marginally lower than the corresponding value of the Shiller estimator (4.94). The OLS estimator, however, had the lowest RMS%E in this case. The in-sample forecast, again, indicated the OLS estimator to have the lowest RMS%E compared to the Almon and Shiller estimators. This conflicts with earlier finding illustrating the superiority of the restricted estimators in a MSE sense. Lastly, the Shiller estimator (1.44) was marginally better than the Almon estimator (1.45) in the in-sample forecast.

On the basis of the RMS%E test, therefore, we find that the unrestricted OLS estimator is the best estimator in terms of forecast accuracy. Since the Almon and Shiller estimators are only marginally different from one another, it seems safer to conclude that the Almon and Shiller estimators perform equally well, but nonetheless, are only next best to the unrestricted OLS estimator.

The contradictory conclusions from the two test procedures warrants further investigation. An ex-

[4] The computed F is identical for the Shiller and Almon estimators since the expression used to compute F is identical regardless of whether the Almon or Shiller estimator is used.

Table 2. Estimated Coefficients of the New York City Transformed Fluid Milk Demand Model: OLS, Almon, and Shiller Estimates, Transformed Data.

	Estimated regression coefficients obtained using		
	Unrestricted estimator	Restricted estimators	
Variables	OLS	Almon	Shiller
Intercept	−2.1740*	−2.2631*	−2.2602*
	(0.4762)@	(0.4881)	(0.4712)
Price of milk	−0.0954*	−0.0961*	−0.0966*
	(0.0484)	(0.0502)	(0.0484)
Price of cola	0.1153*	0.1207*	0.1210*
	(0.0371)	(0.0383)	(0.0370)
Price of coffee	0.0412*	0.0409*	0.0409*
	(0.0083)	(0.0086)	(0.0083)
Income	0.1475	0.1664	0.1657
	(0.1073)	(0.1099)	(0.1061)
Trend	−0.0003*	−0.0003*	−0.0003*
	(0.0001)	(0.0001)	(0.0001)
Dummy variable for:			
January	0.0011	0.0015	0.0012
	(0.0057)	(0.0059)	(0.0057)
February	−0.0060	−0.0046	−0.0054
	(0.0067)	(0.0067)	(0.0066)
March	0.0073	0.0069	0.0070
	(0.0068)	(0.0069)	(0.0068)
April	−0.0119	−0.0137	−0.0130
	(0.0071)	(0.0073)	(0.0071)
May	−0.0219*	−0.0221*	−0.0224*
	(0.0070)	(0.0071)	(0.0069)
June	−0.0378*	−0.0352*	−0.0360*
	(0.0069)	(0.0071)	(0.0069)
July	−0.1002*	−0.1002*	−0.1002*
	(0.0073)	(0.0075)	(0.0072)
August	−0.0925*	−0.0924*	−0.0926*
	(0.0079)	(0.0081)	(0.0078)
September	−0.0229*	−0.0227*	−0.0228*
	(0.0074)	(0.0076)	(0.0074)
October	−0.0009	−0.0010	−0.0010
	(0.0067)	(0.0069)	(0.0067)
November	−0.0130*	−0.0134*	−0.0136*
	(0.0055)	(0.0057)	(0.0055)
Advertising expenditure at time			
t	0.0021*	0.0018*	0.0019*
	(0.0006)	(0.0005)	(0.0005)
t-1	0.0010	0.0019*	0.0016*
	(0.0006)	(0.0003)	(0.0004)
t-2	0.0026*	0.0020*	0.0019*
	(0.0006)	(0.0003)	(0.0004)
t-3	0.0014*	0.0020*	0.0022*
	(0.0006)	(0.0003)	(0.0004)
t-4	0.0034*	0.0020*	0.0021*
	(0.0006)	(0.0003)	(0.0004)
t-5	0.0003	0.0019*	0.0016*
	(0.0006)	(0.0003)	(0.0004)
t-6	0.0025*	0.0018*	0.0020*
	(0.0006)	(0.0005)	(0.0005)
Sum of the advertising coefficients	0.0133	0.0134	0.0133
Adjusted R^2	0.8646	0.8543	0.8519

@: Figures in the parentheses are standard errors.
*: Statistically significant at the 5 percent level.

amination of model disturbances indicated significant first-order autocorrelation in all three models. The Durbin-Watson statistics for the three estimators (shown on the bottom of Table 1) lie below the lower-bound critical value of 1.554 at the 5% level for 200 observations and 20 explanatory variables (Kmenta, p. 764). Given this result, the coefficients obtained earlier, although unbiased and consistent, are not efficient.

To address this problem the original data were transformed using the Prais-Winsten procedure (Park and Mitchell). Results based on the transformed data are qualitatively similar to those obtained earlier except that the estimated income elasticity is much reduced in size (from 0.27 to 0.15) and is no longer significant (table 2). The homogeneity restriction of demand theory, however, is more nearly satisfied by the GLS results.

Re-evaluating the performance of the two estimators, we find that the computed F-statistics are increased in both cases (table 3). In particular, both the Shiller and the Almon estimators yield computed F-values sufficiently large to reject the null hypothesis of SMSE superiority of these estimators against the unrestricted OLS estimator. This confirms our earlier conjecture that the failure to reject the null hypothesis may have been due to the autoregressive disturbances.

Results of the RMS%E test, in this case, are consistent with the above findings. The unrestricted OLS estimator, with the smallest RMS%E, appears to be the best estimator of the coefficients of the transformed model, based on both in-sample and out-of-sample tests. The RMS%Es of the Almon and Shiller estimators were again only marginally different, implying that the Almon and Shiller estimators perform equally well; but again, only next best to the unrestricted OLS estimator.

Based on the foregoing analysis, we conclude that there is no MSE gain from using the Almon or Shiller procedures to estimate the distributed lagged relationship between fluid milk sales and advertising in the New York City Market when the model is corrected for serial correlation. Apparently, the increase in precision offered by the restrictive procedures is insufficient to offset the accompanying bias. This interpretation is consistent with results obtained by Kinnucan (1981) based on data covering a shorter time period. Moreover, the simple correlation coefficients among the lagged advertising variables are all less than 0.39 in absolute value (table 4) indicating that multicollinearity may not be a problem with these data.

The failure to correct for serial correlation can produce misleading results with respect to the presumed superiority of the Almon and Shiller procedures. Note that although the *pattern* of the lagged regression coefficients is significantly affected by the estimating procedure, the *sum* of the coefficients is not. This suggests that if advertising's long-run impact is of primary interest, the biases introduced by the restrictive procedures may be negligible. However, if the pattern of the lagged responses is important (say for advertising pulsing strategies), the Almon and Shiller procedures can produce misleading results.

Concluding Comments

Multicollinearity in models containing distributed lags often results in imprecise estimates of the individual lag parameters. In such cases, the Shiller

Table 3. Performances of the OLS, Almon, and Shiller Estimators Based on the Mean Square Error and Root Mean Square Percent Error Criteria: New York City, Monthly Data, 1971–1990.

	Unrestricted Estimator	Restricted Estimator	
Evaluation Criteria	OLS	Almon	Shiller
Hypothesis Test of MSE Superiority of the Restricted Estimator:			
Computed F:			
Untransformed data	—	2.5555	2.5555
Transformed data	—	4.5962	4.5962
Critical F:	—	*2.9690*	*2.9690*
RMS%E of Out-of-sample prediction:			
Untransformed data	4.8621	4.9328	4.9382
Transformed data	4.9085	5.2116	5.2276
RMS%E of In-sample prediction:			
Untransformed data	1.4046	1.4473	1.4380
Transformed data	1.9739	2.0810	2.0546

Table 4. Simple Correlation Among the Fluid Milk Advertising Lag Variables, New York City Monthly Data, 1971–1988.

Variable	A_t	A_{t-1}	A_{t-2}	A_{t-3}	A_{t-4}	A_{t-5}	A_{t-6}
A_t	1.0000						
A_{t-1}	0.3822	1.0000					
A_{t-2}	0.0003	0.3822	1.0000				
A_{t-3}	−0.0187	0.0005	0.3823	1.0000			
A_{t-4}	−0.0137	−0.0186	0.0005	0.3823	1.0000		
A_{t-5}	−0.0438	−0.0138	−0.0186	0.0006	0.3823	1.0000	
A_{t-6}	0.0650	−0.0440	−0.0138	−0.1821	0.0007	0.3883	1.0000

procedure provides a flexible (*ad hoc*) means of imposing structure on the shape of the lag distribution to obtain more precise estimates. An added advantage of the Shiller procedure is that it subsumes the widely used (and highly restrictive) Almon procedure as a limiting case. However, in applying these procedures, statistical tests such as those suggested by Mittelhammer and Conway should be performed to ensure that the restrictions are compatible with the data. In so doing, the model should be corrected for serial correlation as the presence of serial correlation can lead to erroneous inferences about the performance of the Almon and Shiller procedures. In cases where advertising expenditures vary significantly from period-to-period, as was the case in this study, OLS may still be the best choice for obtaining a quantitative representation of how generic advertisements fade away.

References

Almon, S. "The Distributed Lag Between Capital Appropriations and Expenditures." *Econometrica*, 33(1965):178–196.

Cargill, T., and R.A. Meyer. "Some Time and Frequency Domain Distributed Lag Estimates: A Comparative Monte Carlo Study." *Econometrica*, 42(1974):1031–44.

Chang, H-S., and H.W. Kinnucan. "Advertising, Information, and Product Quality: The Case of Butter." *American Journal of Agricultural Economics*, 73(1991):1195–203.

Fomby, T.B. "MSE Evaluation of Shiller's Smoothness Priors." *International Economic Review*, 20(1979):203–15.

Griliches, Z., G.S. Maddala, R.E. Lucas, and N. Wallace. "Notes on Estimated Aggregate Quarterly Consumption Functions." *Econometrica*, 30(1962):491–500.

Hoerl, A.E., and R.W. Kennard, "Ridge Regression: Biased Estimation for Non-Orthogonal Problems." *Technometrics* 12(1970):55–67.

Judge, G., W.E. Griffins, R.C. Hill, H. Lutkepohl, and T.C. Lee. *The Theory and Practice of Econometrics*, 2nd ed. New York: John Wiley & Sons, 1985.

Kaiser, H.M., D.J. Liu, T.D. Mount, and O.D. Forker. "Impacts of Dairy Promotion from Consumer Demand to Farm Supply." in *Commodity Advertising and Promotion*, eds. H.W. Kinnucan, S.R. Thompson, and H-S. Chang. Ames: Iowa State University Press, 1992.

Kinnucan, H.W. "Demographic Versus Media Advertising Effects on Milk Demand: The Case of the New York City Market." *Northeastern Journal of Agricultural and Resource Economics*, 15(1986):66–74.

Kinnucan, H.W. *Performance of Shiller Lag Estimators: Some Additional Evidence*. A. E. Res. 81-8, Cornell University, 1981.

Kinnucan, H.W., and Forker, O.D. "Seasonality in the Consumer Response to Milk Advertising with Implications for Milk Promotion Policy." *American Journal of Agricultural Economics*, 68(1986):562–71.

Kinnucan, H.W., Chang, H-S., and Venkateswaran, M. "Generic Advertising Wearout." *Review of Marketing and Agricultural Economics*, 61(1993): Forthcoming.

Kmenta, J. *Elements of Econometrics*. (2nd edition). New York: Macmillan Publishing Co., 1986.

Lindley, D.V., and A.F.M. Smith. "Bayes Estimators for the Linear Model." *Journal of the Royal Statistical Society*, B Series, 34(1972):1–41.

Liu, D.J. and Forker, O.D. "Generic Fluid Milk Advertising, Demand Expansion, and Supply Response: The Case of New York City." *American Journal of Agricultural Economics*, 70(1988):229–36.

Maddala, G.S. *Econometrics*, New York: McGraw Hill, 1977.

Mittelhammer, R.C. and Conway, R.K. "Applying Mixed Estimation in Econometric Research." *American Journal of Agricultural Economics*, 70(1988):859–66.

Nerlove, M. and F.V. Waugh. "Advertising Without Supply Control: Some Implications of a Study of the Advertising on Oranges." *Journal of Farm Economics*, 43(1961):813–37.

Park, R.W. and Mitchell, B.M. "Estimating the Autocorrelated Error Model with Trended Data." *Journal of Econometrics* 13(1980):185–201.

Shiller, Robert J. "A Distributed Lag Estimator Derived from Smoothness Priors." *Econometrica*, 41(1973):775–88.

Simon, J.L. and J. Arndt. "The Shape of the Advertising Response Function." *Journal of Advertising Research*. 20(1980):11–28.

Swamy, P.A.V.P. and J.S. Mehta. "On Theil's Mixed Regression Estimator." *Journal of the American Statistical Association*, 64(1969):273–76.

Taylor, W.E. "Smoothness Priors and Stochastic Prior Restrictions in Distributed Lag Estimation." *International Economic Review,* 15(1974):803–04.

Theil, H. *Introduction to Econometrics.* Englewood Cliffs NJ: Prentice Hall, 1978.

Theil, H. and A.S. Goldberger. "On Pure and Mixed Statistical Estimation in Economics." *International Economic Review,* 2(1961):65–78.

Thompson, S.R. and D.A. Eiler. "Producer Returns from Increased Milk Advertising." *American Journal Agricultural Economics,* 57(1975):505–08.

Venkateswaran, M. and H.W. Kinnucan. "Evaluating Fluid Milk Advertising in Ontario: The Importance of Functional Form." *Canadian Journal of Agricultural Economics* 38(1990):471–88.

Wallace, T.D. and C.E. Toro-Vizcarrondo. "Tables for the Mean Square Error Test for Exact Linear Restrictions in Regression." *Journal of the American Statistical Association,* 64(1969):1649–63.

Ward, R.W. and B.L. Dixon. "Effectiveness of Milk Advertising Since the Dairy and Tobacco Adjustment Act of 1983." *American Journal of Agricultural Economics,* 71(1989):730–40.

Waugh, F.V. "Needed Research on the Effectiveness of Farm Products Promotions." *Journal of Farm Economics,* 41(1959):364–76.