Harsharanjeet S. Jagpal
Rutgers University

Ephraim F. Sudit
Rutgers University

Hrishikesh D. Vinod
Bell Laboratories

Measuring Dynamic Marketing Mix Interactions Using Translog Functions

This paper proposes translog sales functions as a more general method for econometric estimation of marketing mix interactions than conventional specifications. A reexamination of the monthly Lydia Pinkham advertising-sales data shows that: (1) the marginal elasticities and productivities of current and lagged advertising depend on the time path of spending; (2) the hypothesis of diminishing returns to advertising is supported after allowing for a general pattern of carryover effects; (3) translog models can be fitted to historical data to measure the effectiveness of advertising policy for control purposes when test market experiments are too expensive, risky, or time-consuming.

Jagpal et al. (1979) showed, using multiplicative nonhomogeneous (MNH) sales response functions, that current econometric specifications for measuring marketing mix interactions and/or carryover effects may be structurally restrictive, regardless of their predictive usefulness. They concluded that there is a need for more flexible specifications of the sales response function which provide greater consistency with marketing theory and hence allow more effective control of marketing instruments. This paper proposes the transcendental logarithmic (translog) formulation because it is more general than the MNH function. First, we develop the rationale for the translog specification. In the second part of the paper, we estimate a translog advertising-sales model using the monthly Lydia Pinkham data. The translog results are shown to be consistent with marketing theory, to differ significantly from previous analyses of the same data, and to have useful policy implications.

Translog Sales Models

Most empirical studies of the sales response function use linear (i.e., constant marginal pro-

(*Journal of Business*, 1982, vol. 55, no. 3)

0021-9398/82/5503-0005$01.50

ductivity) or log-linear (i.e., constant elasticity) specifications (see Clarke [1976] and Weiss, Houston, and Windal [1978] for a review). Recently, Blattberg and Jeuland (1980) have shown, using a micromodeling approach, that several aggregate nonlinear distributed lag advertising-sales model can be deduced depending on the precise behavioral assumptions regarding an individual consumer's exposure to advertising and learning. Clearly, Blattberg and Jeuland's work is a significant methodological contribution. However, from a policy viewpoint, it is important to note two problems with the class of distributed lag and mixed autoregressive/transfer function models of the advertising-sales relationship. First, advertising can have a threshold effect if buyers are unaware of a brand or unwilling to purchase it until they have received several messages (see Krugman [1977] for empirical evidence). Hence, there may be increasing returns at low levels of advertising rather than constant or diminishing returns for *all* levels of spending as implied by the Blattberg-Jeuland model. Second, exposure to advertising messages by the firm can lead to word-of-mouth advertising. In particular, there will be increasing returns to advertising if the proportion of buyers increases at an increasing rate in response to the initial advertising messages (Stigler 1964). Hence the marginal effectiveness of advertising may depend on the time path of advertising outlays. More generally, when several marketing mix variables (e.g., price and advertising) are included simultaneously in the sales response function, theory suggests that the policy variables will interact both contemporaneously and over time (see Parsons and Schultz [1976], pp. 156–64, for a discussion).

Clearly, it is desirable to have a well-developed dynamic theory of marketing incorporating multiple marketing instruments, from which reduced-form specifications of the sales function can be deduced. However, in the absence of such an integrated theory and the practical necessity of estimating sales functions from historical data, it is necessary to choose plausible functional forms for econometric estimation of sales response functions. As a minimum, from a control viewpoint, the specifications chosen should allow *both* the marginal elasticities and productivities of a given policy instrument to depend on the time path of the marketing mix. Further, the functional forms should not require constant or diminishing returns. We propose the translog specification because it provides these features.

The translog functional form, popularized by Griliches and Ringstad (1971); Sargan (1971); and Christensen, Jorgenson, and Lau (1973), is a quadratic approximation to *any* continuous function. Hence it allows for most types of joint effect. Let S denote current sales and X_i the level of the ith marketing policy instrument, current or lagged ($i = 1, \ldots, n$). Then the translog specification of the sales response function $S = f(X_1, \ldots, X_n)$ is (see Christensen et al. 1973, pp. 28–29, pp.

33–34):

$$\ln S = \alpha_0 + \sum_i \alpha_i \ln X_i + \sum_{\substack{i,j \\ j \neq i}} \alpha_{ij} \ln X_i \ln X_j + \sum_i \alpha_{ii} (\ln X_i)^2, \ (i,j = 1, \ldots, n), \tag{1}$$

where the α's are parameters. Hence, the marginal sales elasticity with respect to the ith marketing mix instrument is

$$\epsilon_i = \partial(\ln S)/\partial(\ln X_i) = \alpha_i + \sum_{\substack{j \\ j \neq i}} \alpha_{ij} \ln X_j + 2\alpha_{ii} \ln X_i \ (i,j = 1, \ldots, n) \tag{2}$$

and its marginal sales productivity

$$MP_i = \epsilon_i(S/X_i) \qquad (i,j = 1, \ldots, n), \tag{3}$$

Equations (2) and (3) show that the translog specification, unlike conventional functional forms, explicitly allows both the marginal elasticities *and* productivities of a given marketing instrument to depend on the time path of the entire marketing mix. Further, general types of marketing mix interaction are allowed because equations (2) and (3) are highly nonlinear in their arguments. An additional advantage is that the multiplicative nonhomogeneous (MNH) and Cobb-Douglas (CD) functions are included as special cases. The MNH function is obtained by setting $\alpha_{ii} = 0$ for all i in equation (1) and the CD function by setting $\alpha_{ii} = \alpha_{ij} = 0$ for all i, j. (Because of a printing error, the MNH functional form is incorrectly stated in Jagpal et al. [1979], eq. [3].)

An Empirical Application: The Lydia Pinkham Company

We now illustrate the estimation of translog sales models using the Lydia Pinkham advertising-sales data (see Palda [1963], pp. 32–33, for the data). Monthly rather than annual data are used because (1) attention can be focused on measuring how the marginal productivity (elasticity) of advertising varies with the time path of advertising outlays; (2) the marketing mix (except for advertising) is known to have been relatively stable over this period (January 1954–June 1960). Hence the likelihood of serious specification error resulting from omitted marketing variables is minimal; (3) the assumption that advertising is exogenous may not be valid when annual data is used (see Schmalansee 1972, pp. 10–11); and (4) several previous studies (Palda 1963; Clarke and McCann 1973, 1977; Winer 1979; Jagpal et al. 1979) have examined the monthly series. Hence model comparisons can be made.

Let the advertising-sales relationship be $S_t = f(A_t, A_{t-1}, \ldots, A_{t-l})$ where S_t and A_t denote current sales and advertising, respectively, and advertising has a maximum carryover effect of l periods. From equation (1), the translog sales specification is:

$$\ln S_t = \beta_0' + \sum_i \beta_i \ln A_{t-i} + \sum_{\substack{i,j \\ i \neq j}} \beta_{ij} \ln A_{t-i} \ln A_{t-j} + \sum_i \beta_{ii} (\ln A_{t-i})^2, \qquad (i, j = 0, 1, \ldots, l), \tag{4}$$

where the β's are parameters. Hence the marginal elasticity of current sales with respect to advertising i periods ago is

$$\epsilon_i = \frac{\partial(\ln S_t)}{\partial(\ln A_{t-1})} = \beta_i + \sum_{\substack{j \\ j \neq i}} \beta_{ij} \ln A_{t-j} + 2\beta_{ii} \ln A_{t-i} \qquad (i, j = 0, 1, \ldots, l) \tag{5}$$

and its marginal sales productivity is

$$MP_i = \epsilon_i S_t / A_{t-j} \qquad (i, j = 0, 1, \ldots, l). \tag{6}$$

Note from equations (5) and (6) that both the marginal elasticities and productivities of current and lagged advertising depend nonlinearly on the time path of advertising spending. Further, the translog specification does not force constant or diminishing returns to advertising. Hence it is possible to test for a general pattern of current and lagged advertising effects.

The Empirical Results

Ordinary Least Squares (OLS)

Translog sales functions (see eq. [4]) were estimated for maximum lags of l months ($l \geq 1$). As a first step, the models were screened using both the informal $\bar{R}^2$ and the formal C_p statistic[1] (see Mallows 1973). In general, models with lower C_p are preferred on grounds of parsimony (see Gaver and Geisel [1974], pp. 59–62, for a good discussion of the model simplification problem).

Denoting $\bar{R}^2$ for models with an l-month lag by $\bar{R}_l^2$, the results are: $\bar{R}_1^2 = .3324$, $\bar{R}_2^2 = .4437$, $\bar{R}_3^2 = .5475$, and $\bar{R}_4^2 = .5373$. Similarly, denoting

1. Consider the general linear model $y = X\beta + u$ where $E(u) = 0$ and $E(u'u) = \sigma^2 I$. The C_p statistic is an estimate of the expected sum of squared errors (scaled by $1/\sigma^2$) associated with using a particular estimator to measure the conditional expectation $X\beta$ of y. It is defined by $C_p = RSS_p/\hat{\sigma}^2 - N + 2p$, where p is the number of explanatory variables included, RSS_p is the error sum of squares for a particular p variable regression, N is the sample size, and $\hat{\sigma}^2$ is the estimate of the disturbance variance obtained when all variables are included.

the C_p value for models of lag l compared to the four-period model by $C_p^{(l)}$, the results are: $C_p^{(1)} = 38.24$, $C_p^{(2)} = 25.22$, $C_p^{(3)} = 17.32$, and $C_p^{(4)} = 19$. Based on these results, the two- and three-period models were judged to deserve further analysis.

Tests were then conducted on the estimated residuals for the two- and three-period models to check for model adequacy and specification error. The cumulative periodogram test (see Box and Jenkins 1976, pp. 294–98) was chosen since a monthly time series was being examined and seasonality might easily have led to residuals with a periodic component. The null hypothesis of white noise residuals could not be rejected for either model at the 5% significance level using the Kolmogorov-Smirnoff test (see fig. 1). The rest of the analysis is therefore focused on the remaining problem of multicollinearity in the data.

The degree of multicollinearity in both models is substantial since many eigenvalues λ_i of the correlation matrix are very small.[2] In this case, $\Sigma_i\ 1/\lambda_i = 4{,}817$ and 11,066 for the two- and three-period lag models, respectively, indicating that the OLS estimates have a very high mean squared error in both cases. As a result, many of the t-statistics have very small absolute values (see table 1). Furthermore, many of the computed sales productivities (not reported here) are unstable and have the wrong sign. It is therefore necessary to "treat" multicollinearity. We use the ridge regression method (see App. and Vinod [1978] for a review of multicollinearity).

Ridge Results

The ridge regression results can be examined using the ridge trace in figure 2. Since there are nine regressors in the two-period model, the ridge parameter m lies in the range $0 \leqslant m \leqslant 9$ (see App., eq. [A3]). Figure 2 clearly shows that near the OLS solution at $m = 0$, the standardized b_i^0 are so highly negatively correlated that some pairs of coefficients are estimated to be far apart, even though their pairwise differences may not be statistically significant. As m increases the relative magnitudes and signs of b_i^0 seem to become more stable.

Note that[3] for $l = 2$, six eigenvalues are smaller than $1/p = 0.111$, suggesting that the variance ($= \sigma^2/\lambda_i$) along the last six principal directions is too large. In addition the m-trace becomes stable near the $m = 6$ solution, showing that the $m = 6$ solution is insensitive to small data

2. For the two-period model, the eigenvalues in descending order of magnitude are $\lambda_1 = 5.5505$, $\lambda_2 = 2.8186$, $\lambda_3 = 0.6512$, $\lambda_4 = 0.0118$, $\lambda_5 = 0.0083$, $\lambda_6 = 0.0027$, $\lambda_7 = 0.0014$, $\lambda_8 = 0.0009$, and $\lambda_9 = 0.0004$. Hence $\Sigma_{i=1}^{9}\ 1/\lambda_i = 4{,}817 \gg 9$ for orthogonal data. In the three-period model, the eigenvalues are $\lambda_1 = 6.4771$, $\lambda_2 = 5.0626$, $\lambda_3 = 1.7429$, $\lambda_4 = 0.6658$, $\lambda_5 = 0.0194$, $\lambda_6 = 0.0162$, $\lambda_7 = 0.0061$, $\lambda_8 = 0.0043$, $\lambda_9 = 0.0024$, $\lambda_{10} = 0.0014$, $\lambda_{11} = 0.0008$, $\lambda_{12} = 0.0004$, $\lambda_{13} = 0.0004$, and $\lambda_{14} = 0.0003$. Hence $\Sigma_{i=1}^{14}\ 1/\lambda_i = 11{,}066 \gg 14$ for orthogonal data.

3. See n. 2.

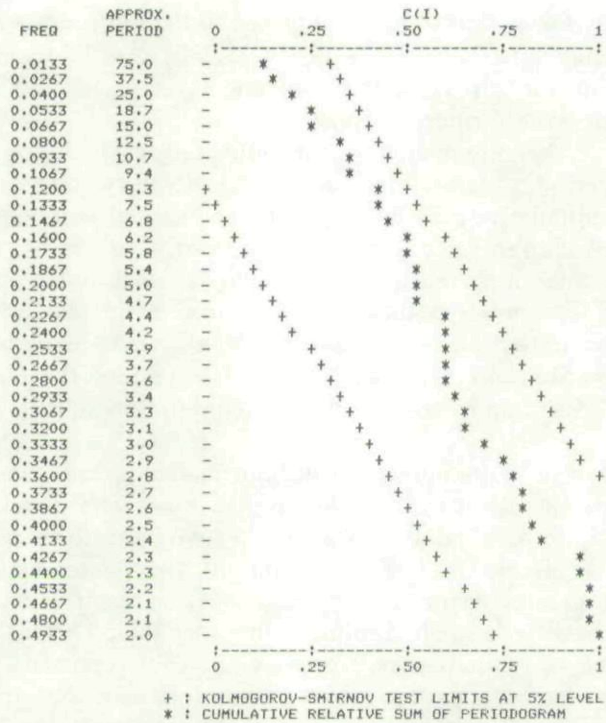

FIG. 1.—Cumulative periodogram of OLS residuals from the two-period translog model.

perturbations (see Vinod 1978, p. 123). Further, $m = 6$ corresponds to the smallest value of the ridge parameter for which the weakly specified quasi-Bayesian prior that the productivities are nonnegative is satisfied.[4] Before choosing the $m = 6$ solution, we tested for model stability by omitting the latest 1, 2, . . . , 10 observations from the data set and reestimating the function in each case. The estimated marginal productivities show remarkable stability near the $m = 6$ solution.[5] Note, however, that on the basis of classical mean squared error (MSE) there is no guarantee that the MSE of the $m = 6$ ridge solution is lower than the MSE at the OLS ($m = 0$) solution. Numerous simulations reviewed in Vinod (1978), however, show that ridge solutions are likely to provide lower MSE in most practical situations.

Further evidence of the preferability of ridge regression in our con-

4. The label quasi-Bayesian is used since from a strict Bayesian viewpoint it is necessary to specify numerical values for the regression parameters. Note that even specific prior numerical knowledge of advertising productivities does not translate into a unique prior for the translog parameters because of the presence of squared and cross-product terms.

5. The detailed tables are available from the authors on request.

TABLE 1 **OLS and Ridge Regression Results for Translog Sales Functions with Maximum Lags of Period l Months**

	$l = 2$		$l = 3$	
Variable	$m = 0$ (OLS)	$m = 6$ (Ridge)	$m = 0$ (OLS)	$m = 10$ (Ridge)
Constant	11.900	6.8491	13.427	6.957
$\ln A_t$	$-.4202$	$-.0055$	$-.3749$	.0035
	(-1.4555)		$(-.8854)$	
$\ln A_{t-1}$	$-.9415$	$-.0006$	$-.7837$	.0070
	(-3.1586)		(-2.613)	
$\ln A_{t-2}$	$-.5334$	$-.0106$	$-.8002$	$-.0008$
	(-1.7597)		(-2.5454)	
$\ln A_{t-3}$	. . .	. . .	$-.5230$	$-.0089$
			(-1.4640)	
$(\ln A_t)(\ln A_{t-1})$	.0234	.0031	.02956	.0019
	(.9737)		(1.2457)	
$(\ln A_t)(\ln A_{t-2})$	$-.0338$	.0001	$-.0517$	.0009
	(-1.1008)		(-1.4767)	
$(\ln A_t)(\ln A_{t-3})$	. . .	. . .	.01781	$-.0021$
			(.5608)	. . .
$(\ln A_{t-1})(\ln A_{t-2})$	.0214	.0014	.0593	.0010
	(.6805)		(1.4757)	
$(\ln A_{t-1})(\ln A_{t-3})$	. . .	. . .	$-.0593$	$-.0001$
			(-1.8504)	
$(\ln A_{t-2})(\ln A_{t-3})$	. . .	. . .	.0359	.0002
			(1.2714)	
$(\ln A_t)^2$	.0456	.0010	.0382	.0009
	(1.9946)		(1.5097)	
$(\ln A_{t-1})^2$	.0645	.0021	.0599	.0014
	(2.3207)		(2.1553)	
$(\ln A_{t-2})^2$	.0559	.0011	.0510	.0007
	(2.0989)		(1.7906)	
$(\ln A_{t-3})^2$	. . .	. . .	.0492	$-.00002$
			(1.9313)	
R^2	.5105	.2350	.6309	.2471
Residual sum of squares (RSS)	.4895	.7532	.3691	.7529
$\hat{\sigma}^2$ (RSS/$N - p - 1$)	.0074	.0111	.0061	.0125
Durbin-Watson statistic	1.66	. . .	1.75	. . .

NOTE.—Figures in parentheses are t-statistics. Note that t-statistics are not defined for ridge estimates because they are biased.

text is based on the following analysis. Since the translog function is an approximation to the true function, and since ridge regression produces biased estimates, it is useful to know how far off one might be by using a biased technique to estimate the translog approximation. For this purpose, the following simulation was done. Assume that the true advertising-sales function is the quartic function

$$\ln S_t = 10 + \ln A_t + \ln A_{t-1} + 0.1 \ln A_t \ln A_{t-1} + 0.1(\ln A_t \ln A_{t-1})^2 + u,$$

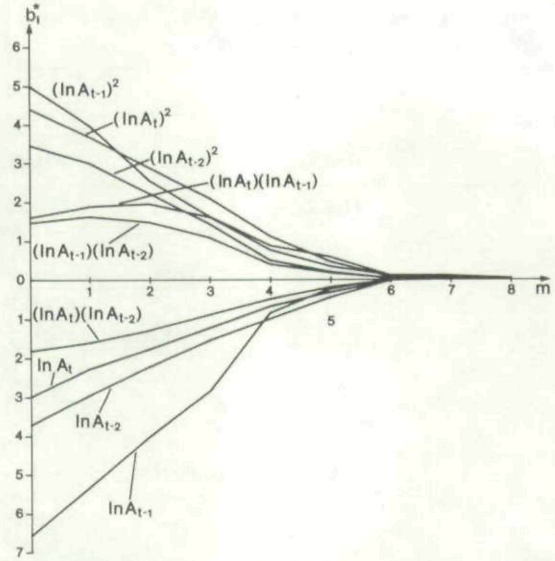

FIG. 2.—Ridge trace on the m-scale for translog sales function with a two-period lag structure.

where u is normally distributed with mean zero and variance one. To construct synthetic sales-advertising data, 29 observations of A_{t-1} and A_t were generated using 30 ordered values of a uniform random variable ranging from 0 to 1.5. For each choice of A_t and A_{t-1}, 20 observations on $\ln S_t$ were generated by adding a randomly drawn value of u to $[10 + \ln A_t + \ln A_{t-1} + 0.1 \ln A_t \ln A_{t-1} + 0.1 (\ln A_t \ln A_{t-1})^2]$. Thus, 20 sets of advertising-sales data were generated, each with 29 observations.

To allow for the fact that the investigator does not know the true form of the model, we specify the translog functional form (see eq. [4], $l = 1$) as an approximation to the true function. The coefficients of the translog specification were estimated for each of the 20 regression problems by OLS as well as by our ridge regression method. The eigenvalues of the correlation matrix of regressors were: 4.759925, 0.2070787, 0.0312787, 0.0017011, and 0.0000159. Since three eigenvalues were smaller than $1/p = 0.2$, we chose $m = 3$, implying a value of $k = 0.08$.

Using equation (5), OLS and ridge marginal sales-advertising elasticities were computed on the basis of the respective coefficient estimates for the 20 regression problems. Two sets of marginal sales-advertising elasticity ratios were computed: (a) OLS-estimated elasticities to true elasticities; and (b) ridge-estimated elasticities to true elasticities. The results for each regression were averaged over the sample observations. In 11 cases, both the OLS and ridge estimates had the wrong sign. However, in 38 out of 40 comparisons, the ridge elasticities were found to be closer in magnitude to the true elasticities than the OLS estimates. These results show that, although specification error cannot be avoided in all cases, by using the translog approximation and ridge estimation, we are likely to obtain better structural estimates of advertising-sales elasticities (productivities) than by using OLS.

Since there are 14 regressors in the three-period model, m lies in the range $0 \leq m \leq 14$. The $m = 10$ solution is indicated, since 10 eigenvalues are smaller than $1/p = 0.071$ and the m-trace (not reproduced here) becomes stable for this choice of m. However, the marginal productivities are economically "meaningless" for $m = 10$ and remain so even though m is chosen as high as 13. We conclude that the use of a three-period model leads to overfitting the data because (a) $\hat{\sigma}^2$ is lower for the two-period ridge solution than for the three-period model (see table 1), (b) the marginal sales productivities of current and lagged advertising (not reproduced here) are very similar to those for the two-period ridge solution, and (c) the lagged marginal sales productivities for the third month are negative *for all months* in the sample period.

The estimated marginal sales productivities of advertising at different lags for the $m = 6$ solution in the two-period model are given in table 2. The major conclusions are: (1) the ridge regression results are reasonable since the current and lagged marginal productivities are always nonnegative. Negative carryover effects would imply the unlikely proposition that past advertising "borrows" sales for nondurables (as distinguished from durables) by reducing future market potential; (2) the hypothesis that the effects of advertising depend on the time path of advertising spending is clearly supported; (3) the marginal elasticities[6] and productivities of advertising vary considerably over time; (4) the current marginal productivity of advertising diminishes with the level of current spending and depends on the pattern of past advertising outlays. This supports the hypothesis of diminishing returns to advertising, after allowing for a general pattern of carryover effects.

6. To illustrate, consider the marginal elasticities and productivities for July 1954 and December 1954 (see table 2). The marginal elasticities are readily obtained from the marginal productivities using eq. (3). The marginal productivities for July 1954 are $MP_0(t) = 0.2445$, $MP_1(t) = 1.0341$, and $MP_2(t) = 0.0617$. However, for December 1954, the marginal productivities are $MP_0(t) = 0.2149$, $MP_1(t) = 0.0804$, and $MP_2(t) = 0.0164$. Similarly, the marginal elasticities for July 1954 are $\epsilon_0(t) = 0.0151$, $\epsilon_1(t) = 0.0345$, and $\epsilon_2(t) = 0.0046$. However, the marginal elasticities for December 1954 are $\epsilon_0(t) = 0.0253$, $\epsilon_1(t) = 0.0518$, and $\epsilon_2(t) = 0.0138$. These results show that both the marginal elasticities and productivities (current and lagged) vary considerably over time.

TABLE 2 Marginal Sales Productivities, $MP(t)$, and Total Marginal Yield, $S_c^{(t)}$ of Advertising for the Two-Period Model Using $m = 6$ Ridge Estimates

Time Period	Sales (in 100's)	Advertising (in 100's)	$MP_0(t)$	$MP_1(t)$	$MP_2(t)$	$S_c^{(t)}$
1954:						
March	1,728	982	.0551	.0781	.0215	.1684
April	1,539	919	.0509	.0939	.0174	.1405
May	1,324	87	.3860	.0747	.0194	1.0188
June	1,264	39	.5325	.5711	.0149	1.7110*
July	1,169	72	.2445	1.0341	.0617	1.2794
August	1,479	467	.0656	.8573	.1444	.2958
September	1,631	1,170	.0396	.1864	.1776	.1363
October	1,546	917	.0520	.0781	.0438	.1616
November	1,459	701	.0616	.0932	.0185	.1604
December	1,087	128	.2149	.0804	.0164	.7902
1955:						
January	1,171	1,014	.0280	.4633	.0184	.1321
February	1,406	1,274	.0342	.0800	.1120	.1313
March	1,619	1,388	.0373	.0786	.0241	.1213
April	1,508	1,071	.0447	.0671	.0185	.1453
May	1,521	537	.0837	.0833	.0169	.2342
June	1,341	123	.2661	.1268	.0173	.7474
July	1,247	60	.3813	.4192	.0237	1.4246*
August	1,262	351	.0706	.8770	.0621	.3345
September	1,419	1,061	.0365	.2086	.1663	.1379
October	1,558	791	.0595	.0849	.0553	.1560
November	1,222	138	.2288	.0815	.0165	.6185
December	1,053	77	.2630	.3293	.0149	1.1677

* Denotes months when the company may have overadvertised.

NOTE.—A detailed table is available from the authors on request.

Marketing Policy Implications

The translog results suggest that historical advertising-sales data can be analyzed to measure advertising effectiveness for control purposes. This is a significant advantage given that it may be too expensive, risky, or time-consuming for a company to use controlled experimentation to measure the effects of advertising.

For the Lydia Pinkham Company, the intertemporal effects of advertising coupled with the pronounced variability of sales productivities over the sample period have potentially useful marketing policy implications. It is interesting to assess the total yield (S_c^t) on a marginal dollar of advertising in month t, where S_c^t is defined for the l-period model by,

$$S_c^t = \sum_{i=0}^{l} \frac{\partial S_{t+i}}{\partial A_t}. \tag{7}$$

For the two-period model chosen,

$$\begin{aligned} S_c^t &= \frac{\partial S_t}{\partial A_t} + \frac{\partial S_{t+1}}{\partial A_t} + \frac{\partial S_{t+2}}{A_t} \\ &= MP_0(t) + MP_1(t+1) + MP_2(t+2). \end{aligned} \tag{8}$$

Note that S_c^t is computed in table 2 on the assumption that A_{t+1} and A_{t+2} were at their actual levels. These results show, first, that in general, the total yield on a marginal dollar of advertising in period t is less than \$1.00. This is not surprising, since the company had an average advertising-sales ratio of 50% for the period. Second, it shows that the total yield of a marginal dollar's advertising is typically larger for months where current advertising was low; and, third, that, assuming Pinkham's total expenses exclusive of advertising and capital costs averaged 25% of sales (see Palda 1964, p. 80), the company overadvertised in some months where outlays were high and underadvertised in some months where advertising was almost shut off completely. This conclusion is not surprising since there is little rationale (except possibly cash flow considerations) for the seasonally changing advertising policy adopted by the company, given that the product does not have an inherently cyclical demand (see table 2, col. 1).

Comparison with Previous Studies

It is interesting to compare the translog results with those from previous studies of the Lydia Pinkham data. Consider first those studies which used distributed lags. Palda (1964, p. 36) found that advertising has a maximum carryover effect of 2 months. Clarke and McCann (1973) using spectral methods also concluded that the effect of advertising on sales is confined to less than 3 months following the advertising (p. 138). In addition, both studies found that the marginal productivities of advertising were the greatest one month after the advertising. These results are consistent with those from the translog model (see table 2). Note, however, that because of the time-invariant linear specifications used in the Palda (1964) and Clarke and McCann (1973) studies all marginal productivities were forced to be constant. Since both studies found the cumulative effects of advertising $S_c^t < 1$, the likely implication is that the company should not have advertised at all regardless of production costs. In contrast, the translog results lead to the more plausible conclusion that the company overadvertised in some but not in all months. Finally, consider Winer's (1979) study which estimates a linear distributed lag model with time-varying parameters. The results were economically meaningless (p. 572), leading Winer to conclude that the company probably used incorrect measurement procedures regarding the timing of advertising and sales.

Consider now the results from previous autoregressive specifications. Palda's (1963, p. 37) "best" autoregressive model leads to $S_c^t = 1.11$.[7] The implication is that, given Palda's assumption that production costs were 25% of sales, the company overadvertised in all months. Clarke and McCann's 1977 study (p. 342) based on the nonlinear least-squares estimation of Houston and Weiss's "current effects" model (1975, p. 473) leads to $S_c^t = 0.274$ and 0.638, depending on the serial correlation model used.[8] These results also imply that the company probably overadvertised in all months. Winer's linear autoregressive model with time-varying parameters, like his distributed lag model, led to economically meaningless results.

Finally, consider the results obtained by Jagpal et al. (1979), who estimate a multiplicative nonhomogeneous sales function using OLS. Interestingly, both the OLS MNH and ridge translog models suggest that the company overadvertised in some but not all months. However, the MNH model leads to a three-period lag structure for advertising, whereas the translog specification leads to a 2-month lag. Further, some lagged marginal elasticities/productivities are negative for the MNH model (see Jagpal et al. 1979, p. 46) whereas all current and lagged elasticities/productivities are positive for the ridge translog model. We believe the ridge translog results to be more realistic than those from the MNH model for two reasons: (1) the translog specification is more general; and (2) advertising is unlikely to have a negative effect on the sales of a frequently purchased product.

7. Palda's best autoregressive model is $S_t = 396 + .39\, S_{t-1} + .27\, A_t + 0.21\, A_{t-1} + 0.16\, A_{t-2}$. This leads to $S_c^t = \Sigma_{k=0}^{\infty}\, \partial S_{t+k}/\partial A_t = 1.111$ by recursive substitution.

8. The two equations are $S_t = \text{constant} + 0.439\, S_{t-1} + 0.154\, A_t$ and $S_t = \text{constant} + 0.671\, S_{t-1} + 0.218\, A_t$, implying $S_c^t = 0.274$ and 0.638, respectively, where $S_c^t = \Sigma_{k=0}^{\infty}\, \partial S_{t+k}/\partial A_t$.

Concluding Remarks

This paper proposes the translog model as a general method for measuring marketing mix interactions and carryover effects. A reanalysis of the monthly Lydia Pinkham advertising-sales data shows that (*a*) the marginal elasticities and productivities of current and lagged advertising depend on the time path of advertising and vary considerably over time; (*b*) there are diminishing returns to current advertising after allowing for a general pattern of lagged effects; and (*c*) the company overadvertised in some but not all months. From a methodological viewpoint, the translog results suggest that historical data can be analyzed to obtain meaningful measures of the effectiveness of marketing policy for control purposes. This is a significant managerial advantage since it can sometimes be too costly, risky, or time-consuming for the marketing manager to use controlled experimentation to measure the effectiveness of marketing policy. Finally, because of its generality, the translog specification can lead to some practical difficulties of estimation if the sales function simultaneously includes several marketing mix variables and several lag periods. In such cases, it may become necessary to make a tradeoff between the generality of the translog specification and the parsimony of less general structural models. One possibility is to fit MNH models to avoid the degrees of freedom problem, making sure to perform the usual econometric tests for model adequacy and specification error.

Appendix
A Review of Ridge Regression

This material is based on Vinod (1976). The reader is referred to the original source or a recent book, Vinod and Ullah (1981), for details.

It is convenient to redefine equation (4) in terms of standardized variables,

$$y = X\beta + u, \tag{A1}$$

where y is a $(T \times 1)$ vector of standardized values of $\ln S_t (1'y = 0)$ where 1 is a vector of ones, X is a $(T \times p)$ matrix of standardized regressors, and u is a $(T \times 1)$ vector of errors satisfying $E(u) = 0$ and $E(uu') = \sigma^2 I$.

Hoerl and Kennard (1970*b*) suggest the ordinary ridge regression (RR) estimator to lower mean squared error:

$$b^* = (X'X + kI)^{-1} X'y, \qquad k > 0. \tag{A2}$$

Note that b^* reduces to the OLS estimator $b^0 = (X'X)^{-1}X'y$ when $k = 0$. Furthermore, $b^* \to 0$, the null vector, as $k \to \infty$.

The empirical problem in RR is that of choosing an appropriate k. We choose k in terms of a new scaling on the horizontal axis called a "multicollinearity allowance" m given by (see Vinod 1976, p. 837):

$$m = p - \sum_i \lambda_i/(\lambda_i + k), \qquad i = 1, \ldots, p, \tag{A3}$$

where the λ_i denote the eigenvalues of $X'X$.

It is readily seen that $m = 0$ corresponds to OLS estimation and that m, unlike k, is bounded from above ($0 \leq m \leq p$).

The m-scale provides several important empirical and theoretical advantages over the k-scale: the ridge trace does not give the spurious appearance of greater stability at larger m; m has a clear interpretation as the assigned deficiency in the rank of $X'X$; and the m-trace unlike the k-trace is easily applied to the generalized RR case.

References

Blattberg, R. C., and Jeuland, A. P. 1980. Micromodeling of the advertising-sales relationship. *Annual Proceedings of the American Marketing Association,* Chicago, pp. 302–6.

Box, G. E. P., and Jenkins, J. M. 1976. *Time Series Analysis: Forecasting and Control.* San Francisco: Holden-Day.

Christensen, L. R.; Jorgenson, D. W.; and Lau, L. J. 1971. Conjugate duality and the transcendental logarithmic production function. *Econometrica* 39 (July): 255–56.

Christensen, L. R.; Jorgenson, D. W.; and Lau, L. J. 1973. Transcendental logarithmic production frontiers. *Review of Economics and Statistics* 55 (February): 28–45.

Clarke, D. G. 1976. Econometric measurement of the duration of advertising effect on sales. *Journal of Marketing Research* 13 (November): 345–57.

Clarke, D. G., and McCann, J. M. 1973. Measuring the cumulative effects of advertising: a reappraisal. *Proceedings of the American Marketing Association* (1973), pp. 135–39.

Clarke, D. G., and McCann, J. M. 1977. Cumulative advertising effects: the role of serial correlation: a reply. *Decision Sciences* 8 (1977): 336–43.

Gaver, K. M., and Geisel, M. S. 1974. Discriminating among alternative models: Bayesian and non-Bayesian methods. In P. Zarembka (ed.), *Frontiers in Econometrics.* New York: Academic Press.

Griliches, Zvi, and Ringstad, V. 1971. *Economies of Scale and the Form of the Production Function.* Amsterdam: North-Holland.

Hoerl, A. E., and Kennard, R. W. 1970*a*. Ridge regression: applications to nonorthogonal problems. *Technometrics* 12 (February): 69–82.

Hoerl, A. E., and Kennard, R. W. 1970*b*. Ridge regression: biased estimation for nonorthogonal problems. *Technometrics* 12 (February): 55–64.

Houston, F. S., and Weiss, D. L. 1975. Cumulative advertising effects: the role of serial correlation. *Decision Sciences* 6: 471–81.

Jagpal, H. S.; Sudit, E. F.; and Vinod, H. D. 1979. A model of sales response to advertising interactions. *Journal of Advertising Research* 19 (June): 41–47.

Krugman, H. E. 1977. Memory with recall: exposure without perception. *Journal of Advertising Research* 17 (August): 7–12.

Mallows, C. M. 1973. Some comments on C_p. *Technometrics* 15 (November): 661–75.

Palda, K. S. 1964. *The Measurement of Cumulative Advertising Effects.* Englewood Cliffs, N.J.: Prentice-Hall.

Parsons, Leonard, and Schultz, R. 1976. *Marketing Models and Econometric Research.* Amsterdam: North-Holland.

Sargan, J. D. 1971. Production functions. Pt. 5 of P. R. G. Layard, M. E. Agar, and D. J. Jones (ed.), *Qualified Manpower and Economic Performance.* London: Penguin.

Schmalansee, R. 1972. *The Economics of Advertising.* Amsterdam: North-Holland.

Stigler, G. J. 1964. A theory of oligopoly. *Journal of Political Economy* 72 (February): 44–61.

Vinod, H. D. 1976. Application of new ridge regression methods to a study of Bell system scale economies. *Journal of the American Statistical Association* 71 (December): 835–41.

Vinod, H. D. 1978. A survey of ridge regression and related techniques for improvements over ordinary least squares. *Review of Economics and Statistics* 60 (February): 121–31.

Vinod, H. D., and Ullah, A. 1981. *Recent Advances in Regression Methods*. New York: Dekker.

Weiss, D. L.; Houston, F. S.; and Windal, P. 1978. The periodic pain of Lydia E. Pinkham. *Journal of Business* 51 (February): 91–102.

Winer, R. S. 1979. An analysis of the time-varying effects of advertising: the case of Lydia Pinkham. *Journal of Business* 52 (October): 563–76.