Econometrica, Vol. 48, No. 5 (July, 1980)

ADVERTISING AND AGGREGATE CONSUMPTION: AN ANALYSIS OF CAUSALITY[1]

BY R. ASHLEY, C. W. J. GRANGER, AND R. SCHMALENSEE

This paper is concerned with testing for causation, using the Granger definition, in a bivariate time-series context. It is argued that a sound and natural approach to such tests must rely primarily on the out-of-sample forecasting performance of models relating the original (non-prewhitened) series of interest. A specific technique of this sort is presented and employed to investigate the relation between aggregate advertising and aggregate consumption spending. The null hypothesis that advertising does not cause consumption cannot be rejected, but some evidence suggesting that consumption may cause advertising is presented.

1. INTRODUCTION

THIS PAPER is concerned with two related questions. The first is empirical: do short-run variations in aggregate advertising affect the level of consumption spending?[2] Many studies find that advertising spending varies pro-cyclically.[3] But firms often use sales- or profit-based decision rules in fixing advertising budgets,[4] so that observed correlation might reflect the effect of advertising on consumers' spending decisions, the effect of aggregate demand on firms' advertising decisions, or some combination of both effects. Previous studies of this empirical question, surveyed in Section 2, do not adequately deal with the problem of determining the direction of causation between consumption and advertising.

The second question with which we are concerned is methodological: how should one test hypotheses about causation in a bivariate time series context? Section 3 proposes a natural approach to such tests that is a direct application of the definition of causality introduced by Granger [**8**]. We argue that it is appropriate to use Box-Jenkins [**2**] techniques to pre-whiten the original series of interest and to use cross-correlograms and bivariate modeling of the pre-whitened series to identify models relating the original series. In our view the out-of-sample forecasting performance of the latter models provide the best information bearing on hypotheses about causation.

The data employed in our study of the advertising/consumption question are described in Section 4, and the results of applying our testing procedure are presented in Section 5. Our main findings are briefly summarized in Section 6.

[1] An earlier version if this paper was written while all three authors were at the University of California, San Diego. Financial support was provided by the Academic Senate of that institution and by National Science Foundation Grant SOC76–14326. The authors are indebted to Robert J. Coen of McCann-Erickson, Dee Ellison of the Federal Trade Commission, Joseph Boorstein and Jonathan Goldberg of the Columbia Broadcasting System, and Robert Parker of the U.S. Department of Commerce for assistance in data preparation, and to Christopher A. Sims and two referees for useful comments. Final responsibility for errors and omissions of course remains with the authors.

[2] The techniques we employ in this study are not well-suited to the detection of very long-run effects that advertising might have on spending patterns, via induced cultural change, for instance.

[3] See, for instance, Simon [**16**, pp. 67–74] and the references he cites.

[4] See, for instance, Kotler [**11**, pp. 350–351], Schmalensee [**15**, pp. 17–18], and the references they cite.

2. PREVIOUS STUDIES

Some evidence against the view that variations in aggregate advertising affect aggregate demand is provided by numerous studies of advertising behavior at cyclical turning points; aggregate advertising generally lags the rest of the economy at such points.[5] Turning point studies do not use much of the information in the time series examined, however, and they do not provide formal tests of hypotheses.

Four relatively recent studies have applied statistical techniques to study the relation between advertising and aggregate demand. In the first of these, Verdon, McConnell, and Roesler [**23**] employed the *Printer's Ink* monthly index of advertising spending (hereinafter referred to as PII). They de-trended PII, GNP, and the Federal Reserve index of industrial production, smoothed all three series with a weighted moving average, and examined correlations between the transformed PII series and the other two transformed series at various leads and lags and for various periods. The correlations obtained showed no clear patterns.

In a critique of this study, Ekelund and Gramm [**7**] argued that consumption spending, rather than GNP or the index of industrial production, should be used in tests of this sort. They regressed de-trended quarterly advertising data from Blank [**1**] on de-trended consumption spending, and all regressions were insignificant.

Taylor and Weiserbs [**21**] considered four elaborations of the Houthakker–Taylor [**10**] consumption function that included contemporaneous advertising. Annual data were employed, consumption and income were expressed in 1958 dollars, and advertising spending was used both in current dollars and deflated by the GNP deflator. One of their models performed well, and it had a significant advertising coefficient even when re-estimated by a two-stage least squares procedure that treated advertising as endogenous. Taylor and Weiserbs concluded that aggregate advertising has a significant effect on aggregate consumption.

There are at least four serious problems with this study, however. First, as the authors acknowledge, their conclusion rests on the somewhat restrictive maintained hypothesis that the Houthakker–Taylor framework is correct. Second, the GNP deflator is not a particularly good proxy for the price of advertising messages.[6] Third, their two-stage least squares procedure may not deal adequately with advertising's probable endogeneity. It rests on a rather ad hoc structural equation for advertising spending. Further, all structural equations have lagged endogenous variables, so that the consistency of the estimators depends critically on the disturbances being serially uncorrelated.[7] Fourth, annual

[5] See Simon [**16**, pp. 67–74] and Schmalensee [**15**, pp. 17–18] for surveys of these studies.

[6] Using the sources described in the Appendix, an implicit deflator for the six media considered there was constructed for the period 1950–1975. Over that period, it grew at 2.2 per cent per year, while the GNP deflator increased an average of 3.5 per cent per year. The simple correlation between the first differences of the two series was only .60.

[7] We are told that Durbin's [**6**] test did not reject the null hypothesis of no serial correlation, but that test explicitly considers only the alternative of first-order autoregression. Moreover, the small sample properties of Durbin's test are not well understood [**12**], and Taylor and Weiserbs have only 35 residuals.

data are likely to be innappropriate here. In a survey of econometric studies of the effects of advertising on the demand for individual products, Clarke [**4**] finds that between 95 per cent and 100 per cent of the sales response to a maintained increase in advertising occurs within one year. Similarly, Schmalensee's [**15**, Ch. 3] estimates of aggregate advertising spending functions indicate that between 75 per cent and 85 per cent of the advertising response to a maintained increase in sales occurs within one year. These findings suggest that in this context so much information is lost by aggregation over time that annual data simply cannot contain much information about the direction of causation.

Finally, Schmalensee [**15**, pp. 49–58] employed an extension of Blank's [**1**] quarterly advertising series, deflated to allow for changes in media cost and effectiveness, in connection with several standard aggregate consumption equations specified in constant dollars per capita. Using instrumental variables estimators, the previous quarter's advertising, the current quarter's advertising, and the following quarter's advertising were added one at a time to the consumption equations. It was found that current advertising generally out-performed lagged advertising, and future advertising generally outperformed current advertising in fitting the data. Schmalensee took this pattern to imply that causation ran from consumption to advertising, reasoning that if advertising were causing consumption, past advertising would have outperformed future advertising.

Schmalensee's study has at least two major weaknesses. First, no tests of significance are applied to the observed performance differences. Second, nothing rules out the possibility that advertising is causing consumption as well as being caused by it. If both effects are present, both affect observed performance differentials, and these can in principle go in either direction.

It seems clear that in order to go beyond these studies, one must employ a statistical procedure explicitly designed to test hypotheses about causality in a time-series context. Accordingly, we now present such a procedure.

3. TESTING FOR CAUSALITY

The phrase 'X causes Y' must be handled with considerable delicacy, as the concept of causation is a very subtle and difficult one. A universally acceptable definition of causation may well not be possible, but a definition that seems reasonable to many is the following: Let Ω_n represent all the information available in the universe at time n. Suppose that at time n optimum forecasts are made of X_{n+1} using all of the information in Ω_n and also using all of this information apart from the past and present values Y_{n-j}, $j \geqslant 0$, of the series Y_t. If the first forecast, using all the information, is superior to the second, than the series Y_t has some special information about X_t, not available elsewhere, and Y_t is said to cause X_t.

Before applying this definition, an agreement has to be reached on a criterion to decide if one forecast is superior to another. The usual procedure is to compare the relative sizes of the variances of forecast errors. It is more in keeping with the spirit of the definition, however, to compare the mean-square errors of post-sample forecasts.

To make the suggested definition suitable for practical use a number of simplifications have to be made. Linear forecasts only will be considered, together with the usual least-squares loss function, and the information set Ω_n has to be replaced by the past and present values of some set of time series, $R_n: \{X_{n-j}, Y_{n-j}, Z_{n-j}, \ldots, j \geq 0\}$. Any causation now found will only be relative to R_n and spurious results can occur if some vital series is not in this set.

The simplest case is when R_n consists of just values from the series X_t and Y_t, where now the definition reduces to the following.

Let MSE(X) be the population mean-square of the one-step forecast error of X_{n+1} using the optimum linear forecast based on X_{n-j}, $j \geq 0$, and let MSE(X, Y) be the population mean-square of the one-step forecast error of X_{n+1} using the optimum linear forecast based on X_{n-j}, Y_{n-j}, $j \geq 0$. Then *Y causes X* if MSE(X, Y) < MSE(X).

With a finite data set, some test of significance could be used to test if the two mean-square errors are significantly different; one such test is presented below and employed in Section 5. As the scope of this definition has been greatly circumscribed by the simplifications used, the possibility of incorrect conclusions being reached is expanded,[8] but at least a useable form of the definition has been obtained. This definition of causation (stated in terms of variances rather than mean-square errors) was introduced into the economic literature by Granger [**8**]; it has been applied by Sims [**17**] and numerous subsequent authors employing a variety of techniques. (See [**14**] for a survey.)

The next several paragraphs present the five-step approach to the analysis of causality (as defined above) between a pair of time series X_t and Y_t that is employed in Section 5, below. The remainder of this Section then discusses the rationale for our approach.

(i) Each series is pre-whitened by building single-series ARIMA models using the Box-Jenkins [**2**] procedure. Denote the resulting residuals by εx_t and εy_t.

(ii) Form the cross-correlogram between these two residual series, i.e., compute

$$\rho_k = \text{corr}\ (\varepsilon x_t, \varepsilon y_{t-k})$$

for positive and negative values of k. If any ρ_k for $k > 0$ are significantly different from zero, there is an indication that Y_t may be causing X_t, since the correlogram indicates that past Y_t may be useful in forecasting X_t. Similarly, if any ρ_k is significantly non-zero for $k < 0$, X_t appears to be causing Y_t. If both occur, two-way causality, or feedback, between the series is indicated.

Unfortunately, the sampling distribution of the ρ_k depends on the exact relationship between the series. On the null hypothesis of no relationship, it is well known that the ρ_k are asymptotically distributed as independent normal with means zero and variances $1/n$, where n is the number of observations employed [**9**, p. 238], but the experience shows that the test suggested by this result must be

[8] Sims [**20**] provides a discussion of possible spurious sources of apparent causation in applications of this definition. In Section 6, below, we consider the likely importance of these in our empirical analysis.

used with extreme caution in finite samples.[9] In practice, we also use a priori judgement about the forms of plausible relations between economic time series. Thus, for example, a value of ρ_1 well inside the interval $[-2/\sqrt{n}, +2/\sqrt{n}]$ might be tentatively treated as significant, while a substantially larger value of ρ_7 might be ignored if ρ_5, ρ_6, ρ_8, and ρ_9 are all negligible.

This step is perfectly analogous to the univariate Box-Jenkins identification step, where a tentative specification is obtained by judgmental analysis of a correlogram. The key word is "tentative"; the indicated direction of causation is only tentative at this stage and may be modified or rejected on the basis of subsequent modelling and forecasting results.[10]

(iii) For every indicated causation, a bivariate model relating the residuals is identified, estimated, and diagnostically checked. If only one-way causation is present, the appropriate model is unidirectional and can be identified directly from the shape of the cross-correlogram, at least in theory. However, if the series are related in a feedback fashion, the cross-correlogram has to be unraveled into a pair of transfer functions to help with model identification, by a procedure developed by Granger and Newbold [**9**, Ch. 7].

(iv) From the fitted mode for residuals, after dropping insignificant terms, the corresponding model for the original series is derived, by combining the univariate models with the bivariate model for the residuals. It is then checked for common factors, estimated, and diagnostic checks applied.[11]

(v) Finally, the bivariate model for the original series is used to generate a set of one-step forecasts for a post-sample period. The corresponding errors are then compared to the post-sample one-step forecast errors produced by the univariate model developed in step (i) to see if the bivariate model actually does forecast better.[12] The use of sequential one-step forecasts follows directly from the definition above and avoids the problem of error build-up that would otherwise occur as the forecast horizon is lengthened.

Because of specification and sampling error (and perhaps some structural change) the two forecast error series thus produced are likely to be cross-correlated and autocorrelated and to have non-zero means. In light of these

[9] One must apparently be even more careful with the Box-Pierce [3] test on sums of squared ρ_k; see [**5**].

[10] See Granger and Newbold [**9**, pp. 230–266] for a fuller discussion of this approach. Unpublished simulations performed at UCSD (e.g., C. Chiang, "An Investigation of Relationships Between Price Series," unpublished dissertation, Department of Economics, 1978) find that it rarely signals non-existent causations but lacks power in that subtle causations are not always detected.

[11] OLS estimation suffices to produce unbiased estimates, since all the bivariate models considered are reduced forms. It also allows one to consider variants of one equation without disturbing the forecasting results from the other, and it is computationally simpler. On the other hand, where substantial contemporaneous correlation occurs between the residuals, seemingly-unrelated regressions GLS estimation can be expected to yield noticeably better parameter estimates and post-sample forecasts. All estimation in this study is OLS; a re-estimation of our final bivariate model using GLS might strengthen our conclusions somewhat.

[12] Alternatively, one might fit both models to the sample period, produce forecasts of the first post-sample observation, re-estimate both models with that observation added to the sample, forecast the second post-sample observation, and so on until the end of the post-sample period. This would, of course, be more expensive than the approach in the text.

problems, no direct test for the significance of improvements in mean-squared forecasting error appears to be available. Consequently, we have developed the following indirect procedure.

For some out-of-sample observation, t, let e_{1t} and e_{2t} be the forecast errors made by the univariate and bivariate models, respectively, of some time series. Elementary algebra then yields the following relation among sample statistics for the entire out-of-sample period:

$$\text{(1)} \qquad \mathrm{MSE}(e_1)-\mathrm{MSE}(e_2)=[s^2(e_1)-s^2(e_2)]+[m(e_1)^2-m(e_2)^2],$$

where MSE denotes sample mean-squared error, s^2 denotes sample variance, and m denotes sample mean. Letting

$$\text{(2)} \qquad \Delta_t=e_{1t}-e_{2t}, \qquad \text{and} \quad \Sigma_2=e_{1t}+e_{2t},$$

equation (1) can be re-written as follows, even if e_{1t} and e_{2t} are correlated [**9**, p. 281]:

$$\text{(3)} \qquad \mathrm{MSE}(e_1)-\mathrm{MSE}(e_2)=[\widehat{\mathrm{cov}}\,(\Delta, \Sigma)]+[m(e_1)^2-m(e_2)^2],$$

where $\widehat{\mathrm{cov}}$ denotes the sample covariance over the out-of-sample period.

Let us assume that both error means are positive; the modifications necessary in the other cases should become clear. Consider the analogue of (3) relating population parameters instead of sample statistics, and let cov denote the population covariance and μ denote the population mean. From (3), it is then clear that we can conclude that the bivariate model outperforms the univariate model if we can reject the joint null hypothesis cov $(\Delta, \Sigma)=0$ and $\mu(\Delta)=0$ in favor of the alternative hypothesis that both quantities are nonnegative and at least one is positive.

Now consider the regression equation

$$\text{(4)} \qquad \Delta_t=\beta_1+\beta_2[\Sigma_t-m(\Sigma_t)]+u_t,$$

where u_t is an error term with mean zero that can be treated as independent of Σ_t.[13] From the algebra of regression, the test outlined in the preceding paragraph is equivalent to testing the null hypothesis $\beta_1=\beta_2=0$ against the alternative that both are nonnegative and at least one is positive. If either of the two least squares estimates, $\hat{\beta}_1$ and $\hat{\beta}_2$, is significantly negative, the bivariate model clearly cannot be judged a significant improvement. If one estimate is negative but not significant, a one-tailed t test on the other estimated coefficient can be used.

[13] In fact, this independence assumption must be violated; a bit of algebra shows that in the *population*,

$$\mathrm{cov}\,(\Sigma_t, u_t)=\mathrm{cov}(\Sigma_t, \Delta_t)-\beta_2\,\mathrm{var}\,(\Sigma_t)$$

where var denotes the population variance. On the other hand, it is clear that β_1 is estimated without bias, and it can be shown that the bias in $\hat{\beta}_2$ is equal to the difference between the sample and population values of cov $(\Sigma_t, u_t)/\mathrm{var}\,(\Sigma_t)$. This bias should thus be of negligible importance in moderate samples.

If both estimates are positive, an F test of the null hypothesis that both population values are zero can be employed. But this test is, in essence, four-tailed; it does not take into account the signs of the estimated coefficients. If the estimates were independent, it is clear that the probability of obtaining an F statistic greater than or equal to F_0, say, *and* having both estimates positive is equal to one-fourth the significance level associated with F_0. Consideration of the possible shapes of iso-probability curves for $(\hat{\beta}_1, \hat{\beta}_2)$ under the null hypothesis that both population values are zero establishes that the true significance level is never more than half the probability obtained from tables of the F distribution. If both estimates are positive then, one can perform an F test and report a significance level equal to half that obtained from the tables.

The approach just described differs from others that have been employed to analyze causality in its stress on models relating the original variables and on post-sample forecasting performance. We now discuss these two differences.

Many applications of the causality definition considered here (e.g., [**13**]) essentially stop at our stage (ii) and thus consider only the sample cross-correlogram of the prewhitened series. For a variety of reasons, it seems to us unwise to rest causality conclusions entirely on correlations between estimated residuals. Sims [**19**], for instance, has argued that there may be a tendency for such correlations to be biased toward zero because of specification error. To see the nature of the argument, suppose Y causes X, so that the appropriate model for X is bivariate. Estimation of such a model on the original series would allow the data to indicate the relative importance of "past X" and "past Y" in forecasting X. Prewhitening X, on the other hand, involves use of a misspecified model in this case, since "past Y" should be included. As in standard discussions of omitted variable bias, correlation between "past X" and "past Y" will tend to lead the misspecified univariate model to over-state the importance of "past X" in forecasting current X. The correlation between the residual series from this model and (original or prewhitened) "past Y" will accordingly be biased toward zero.

Thus, models directly relating the original variables provide a sounder, as well as a more natural basis for conclusions about causality. As has been argued in detail by Granger and Newbold [**9**, Sect. 7.6], however, prewhitening and analysis of the cross-correlogram of the prewhitened series are useful steps in the identification of models relating the original series, since the cross-correlogram of the latter is likely to be impossible to interpret sensibly. Because the correlations between the prewhitened series (the ρ_k) have unknown sampling distributions, this analysis involves subjective judgements, as does the identification step in univariate Box-Jenkins analysis. In neither case is an obviously better approach available, and in both cases the tentative conclusions reached are subjected to further tests.

It is somewhat less clear how out-of-sample data are optimally employed in an analysis of causality. This question is closely related to fundamental problems of model evaluation and validation and is complicated by sampling error and possible specification error and time-varying coefficients. An attempt to sort all this out would clearly carry us well beyond the bounds of the present essay.

However, we think the riskiness of basing conclusions about causality entirely on within-sample performance is reasonably clear. Since the basic definition of causality is a statement about forecasting ability, it follows that tests focusing directly on forecasting are most clearly appropriate. Indeed, it can be argued that goodness-of-fit tests (as opposed to tests of forecasting ability) are contrary in spirit to the basic definition.[14] Moreover, within-sample forecast errors have doubtful statistical properties in the present context when the Box-Jenkins methodology is employed. While the power of that methodology has been demonstrated in numerous applications and rationalizes our use of it here, it must be noted that the identification (model specification) procedures in steps (i)–(iv) above involve consideration and evaluation of a wide variety of model formulations. A good deal of sample information is thus employed in specification choice, and there is a sense in which most of the sample's real degrees of freedom are used up in this process. It thus seems both safer and more natural to place considerable weight on out-of-sample forecasting performance.

The approach outlined above uses the post-sample data only in the final step, as a test track over which the univariate and bivariate models are run in order to compare their forecasting abilities. This approach is of course vulnerable to undetected specification error or structural change. Partly as a consequence of this, the likely characteristics of post-sample forecast errors render testing for performance improvement somewhat delicate, as we noted above. Finally, the appropriate division of the total data set into sample and post-sample periods in this approach is unclear. (We say a bit about this in light of our advertising/consumption results in Section 6.) These are nontrivial problems. But at present, we see no way to make more use of the post-sample data that does not encounter apparently equally severe problems.[15]

We do not want to seem overly dogmatic on this issue. Our basic point is simply that model specification (perhaps especially within the Box-Jenkins framework)

[14] If one finds that one model (using a wider information set, say) fits better than another, one is really saying "If I had known that at the beginning of the sample period, I could have used that information to construct better forecasts *during* the sample period." But this is not strictly operational and thus seems somewhat contrary in spirit to the basic definition of causality that we employ.

[15] Two possibilities have been suggested. Both involve goodness-of-fit tests, about which we have some misgivings as footnote 14 indicates. (i) One could use asymptotic variants of covariance analysis ("Chow tests") to investigate the appropriateness of the sample specification for the post-sample period. Assuming this test is passed by both univariate and bivariate models, goodness-of-fit in the pooled sample could be used to compare model performance. However, depending on the sample/post-sample split, final conclusions may be inordinately influenced by the same sample data that guided specification choice. Moreover, it is not clear what should be done if either model fails the stability test. Simply concluding that no inferences about causality can be made seems unsatisfactory, but any other alternative must run the risk of "mining" the post-sample data. Similar problems arise if the post-sample data are used for any critical diagnostic tests on the models selected. (In addition, appropriate testing procedures are unclear, since sampling error implies likely non-whiteness of post-sample errors.) (ii) One could simply re-estimate the univariate and bivariate models derived from the sample using only the post-sample data and compare fits for this period. Depending on the sample/post-sample split, again, these estimates may be unreliable. However, this approach avoids mining the post-sample data, and it yields error series with zero means. But these series will not necessarily be white. Moreover, it seems odd to carry over the specification from the sample period but otherwise to ignore the data on which it is based. Still, if very long time series are available, this second approach may be a viable alternative to the one discussed in the text.

may well be infected by sampling error and polluted by data mining, so that it is unwise to perform tests for causality on the same data set used to select the models to be tested. The procedure outlined above seems to handle this problem sensibly.

4. THE DATA

In light of the evidence on the lengths of the relevant lags noted in Section 2, above, the use of quarterly data seems necessary if defensible judgements are to be reached about the causal relation, if any, between aggregate advertising and aggregate consumption. This section discusses the time series variables used to study that relation. All variables are computed for the period 1956–1975, yielding a total of 80 quarterly observations. A logarithmic transformation of all series is employed to reduce observed heteroscedasticity.

We know of two series of U.S. quarterly advertising spending estimates: the PII and its successors,[16] and extensions by the Columbia Broadcasting System (CBS) of Blank's [**1**] series. The Appendix indicates why we elect to use the CBS figures here and describes their employment in the computation of *ADN*: national advertising in major media, current dollars per capita, seasonally adjusted.

In [**15**, Ch. 3] it is argued that percentage-of-sales decision rules for advertising spending have the strongest theoretical rationale when both advertising and sales are in nominal (current dollar) terms. On the other hand, one might expect the impact of advertising on consumer spending to be most apparent when both quantities are in real terms. Real advertising data are obtained by adjusting expenditure figures to take into account changes in both rates and audience sizes; real advertising per capita must measure the number of messages to which an average person is exposed. There apparently exist no quarterly advertising cost or price indices that could be used directly to obtain real advertising, however. One must either deflate nominal spending totals by some arbitrarily chosen alternative quarterly price indices or use interpolated values of annual advertising price indices. Since the cost of advertising messages has changed relative to prices of other goods and services (see footnote 6, above), it seems safest to interpolate. The Appendix describes the use of interpolated annual indices to calculate *ADR*: national advertising in major media, 1972 dollars per capita, seasonally adjusted.

The following consumption series were based on data from the January and March, 1976 issues fo the *Survey of Current Business*: *CTN*: total personal consumption expenditure, thousands of current dollars per capita, seasonally adjusted; *CGN*: personal consumption expenditure on goods, thousands of current dollars per capita, seasonally adjusted; *CTR*: total personal consumption expenditure, thousands of 1972 dollars per capita, seasonally adjusted; *CGR*: personal consumption expenditure on goods, thousands of 1972 dollars per capita, seasonally adjusted.

[16] These are the *Marketing/Communications* Index and, beginning in 1971, the McCann-Erickson Index. In recent years, all these estimates have been prepared by McCann-Erickson and reported monthly in the *Survey of Current Business.*

The main reason for considering consumption spending on goods only is that the bulk of services consumption is devoted to items that are not heavily nationally advertised, though they may be locally advertised [**15**, pp. 62–64]. Moreover, services consumption is notoriously stable about its trend.

It is relatively well known [**18**, **24**] that the standard methods of seasonal adjustment, which have been applied to the series discussed thus far, can lead to sizeable biases in contexts such as ours.[17] We would have preferred to begin with a set of time series that had not been seasonally adjusted, and some of the results reported below would seem to support this prejudice. Of the series discussed so far, however, it was only possible to obtain unadjusted numbers corresponding to *CTN* and *CGN*. Based on unpublished data supplied by the U.S. Department of Commerce, we assembled *UCTN*: total personal consumption expenditure, thousands of current dollars per capita, not seasonally adjusted; and *UCGN*: personal consumption expenditure on goods, thousands of current dollars per capita, not seasonally adjusted.

All series employed are natural logarithms (as noted above) of quarterly totals at annual rates. All are available from the authors on request.

5. EMPIRICAL RESULTS

We initially considered only the first six (seasonally adjusted) series described in Section 4. It was decided to retain the last 20 observations to evaluate out-of-sample forecasting performance, since we reached the judgement that fewer than 60 data points would not permit adequate identification and estimation in this case.

As per step (i) of the approach outlined in Section 3, univariate time series models were identified and estimated for the six series considered using the sixty quarterly observations from 1956 through 1970.[18] None of the six residual (prewhitened) series showed significant serial correlation.

Proceeding to step (ii), cross-correlograms of the appropriate pairs of residual series were computed. Letting εx_t denote the residual from a univariate model for the variable x_t, this involved computation of corr $(\varepsilon adn_t, \varepsilon ctn_{t-k})$, corr $(\varepsilon adn_t, \varepsilon cgn_{t-k})$, corr $(\varepsilon adr_t, \varepsilon ctr_{t-k})$, and corr $(\varepsilon adr_t, \varepsilon cgr_{t-k})$ for k between -10 and $+10$. All four cross-correlograms were strikingly similar, indicating that it made little difference whether we worked in nominal or real terms, or whether we used total or goods consumption. All four showed a strong contemporaneous correlation ($k=0$), which, however, provides no information on the direction of causation. Sizeable positive correlations for $k=-1$ suggested that advertising might be causing consumption, while similar correlations for $k=+1$, $+2$, and $+3$ suggested consumption causing advertising.

All four of these cross-correlograms showed substantial negative values at

[17] See the Appendix, especially footnote 29. Since the Census X-11 procedure used on these data involves a two-sided filter for most of the sample period, its employment in an investigation of causation is particularly worrisome.

[18] Descriptions of these models and other statistical results not reported here are contained in an earlier version of this essay, available as Discussion Paper 77–9 from the Department of Economics, University of California, San Diego (La Jolla, CA 92093).

$k=+7$ and $k=-5$. Since the neighboring correlations were clearly negligible, we found it difficult to interpret these in causal terms. Suspecting that the correlations at $k=-5$ and, possibly, $k=+7$ were artifacts of the seasonal adjustment procedures applied to the data, we obtained the unadjusted consumption expenditure series $UCTN_t$ and $UCGN_t$ defined above. In light of the discussion of services consumption in Section 4 and the similarity of the cross-correlograms discussed above, it was decided to confine our attention initially to $UCGN_t$, current dollar consumption spending on goods.

Proceeding as before, the following univariate model was identified, estimated, and checked:

$$\text{(C.1)} \qquad (1-B)(1-B^4)UCGN_t = \underset{(.00043)}{.00086} + (1 - \underset{(.082)}{.204}B^2 - \underset{(.075)}{.747}B^4)\varepsilon ugn_t,$$

where B is the lag or backward shift operator, numbers in parentheses are standard errors, and $\varepsilon ucgn_t$ is a residual series, as above. (The presence of $(1-B^4)$ reflects the use of seasonal differencing.) The corresponding univariate model for advertising was the following:

$$\text{(A.1)} \qquad (1-B)ADN_t = \underset{(.0022)}{.00911} + (1 - \underset{(.13)}{.256}B^5)\varepsilon adn_t.$$

The cross-correlogram between the residual series from these models is given as row 1 in Table I. Use of unadjusted consumption substantially reduced the anomalous correlations at $k=-5$ and $k=+7$. (An approximate 95 per cent confidence interval for any single correlation here is $[-.27, +.27]$.) This suggests that these correlations were in fact artifacts of the use of standard seasonal adjustment procedures. In light of these results, it was decided to restrict further attention to the relation between ADN_t and $UCGN_t$.[19] The sample and post-sample performance of the univariate models (A.1) and (C.1) are shown in Table II.

As per Section 3, we now proceed to step (iii), modeling the relation between the univariate residual (i.e., prewhitened) series εadn_t and $\varepsilon ucgn_t$. Examination of row 1 of Table I shows that the contemporaneous $(k=0)$ correlation is large compared to $1/\sqrt{n}$, which is .14 here. The correlation at $k=+1$ is not significant on the usual test, but it and the $k=0$ term together suggest a sensible lag structure that deserves further examination. In contrast, the $k=-1$ and $k=-2$ terms are clearly negligible. The correlations at $k=-3$, -4, and -5 are nonnegligible, but it is hard to put them together with the $k=0$ term (and the negligible terms in between) to form a plausible lag structure. Hence the cross-correlogram tentatively suggests that a unidirectional model, in which $\varepsilon ucgn_t$ causes, but is not caused by, εadn_t is appropriate.

Before proceeding on this assumption, however, it seems appropriate to test it by constructing a forecasting model for $\varepsilon ucgn_t$ employing lagged values of εadn_t. The best model obtained, called (CA.1) in Table II, includes εadn_{t-k} for $k=3, 4,$

[19] Note that this means that, as mentioned in footnotes 17 and 29, the advertising series has been put through a two-sided filter, while the consumption series has not been. In general, one would expect this to bias our results toward a finding that advertising causes consumption, if the series are actually causally related.

TABLE I

AUTO- AND CROSS-CORRELOGRAMS FOR RESIDUAL SERIES

Row	Residual Series	Correlation for $k =$ −7	−6	−5	−4	−3	−2	−1	0	+1	+2	+3	+4	+5	+6	+7
1	εadn_t, $\varepsilon ucgn_{t-k}$	.05	.06	−.14	−.13	−.19	.09	.04	.50	.18	−.02	.16	−.13	.16	−.13	−.13
2	ηadn_t, ηadn_{t-k}	−.13	.09	−.20	.00	.19	−.03	.01	1.0	.01	−.03	.19	.00	−.20	.09	−.13
3	$\varepsilon ucgn'_t$, $\varepsilon ucgn'_{t-k}$	−.15	−.12	.08	−.05	−.09	.01	−.03	1.0	−.03	.01	−.09	−.05	.08	−.12	−.15
4	$\varepsilon ucgn'_t$, ηadn_{t-k}	−.14	−.13	.10	.05	.16	.18	−.10	.50	.13	.05	−.08	−.14	−.09	−.01	−.01

TABLE II

PERFORMANCE OF UNIVARIATE AND BIVARIATE MODELS

Row	Model	Model Type	Error Term	Sample Variance[a]	Post-Sample MSE[b]
1	(A.1)	Univariate	εadn	454	722
2	(AC.1)	Bivariate on Residuals	γadn	435	600
3	(AC.2)	Bivariate on Original Series	ηadn	416	533
4	(C.1)	Univariate	$\varepsilon ucgn$	245	261
5	(CA.1)	Bivariate on Residuals	$\gamma ucgn$	213	290
6	(C.2)	Univariate	$\varepsilon ucgn'$	268	234
7	(CA.2)	Bivariate on Original Series	$\eta ucgn$	263	222

[a] Sample period (1956–70) variance $\times 10^6$; not corrected for degrees of freedom.
[b] Post-sample period (1971–75) mean squared error of one-step-ahead forecasts $\times 10^6$.

and 5 only. A comparison of rows 4 and 5 of Table II shows that this model performs quite badly in the post-sample period. These findings support the tentative identification of unidirectional causation.

Accordingly, we now consider the impact of prewhitened consumption on prewhitened advertising. The form of the cross-correlogram suggests that an appropriate identification for a model of this relationship is

$$(1-\alpha B)\varepsilon adn_t = (\beta_1 + \beta_2 B)\varepsilon ucgn_t + \gamma adn_t.$$

The αB term is included because it is necessary to have polynomials in the lag operator, B, of the same order on both sides of the equation since the model represents a unidirectional relationship between two white noise series [**9**, Ch. 7]. If a purely forecasting model is constructed using this identification (by omitting the contemporaneous term), one obtains

$$\text{(AC.1)} \quad (1+\underset{(.15)}{.200}B)\varepsilon adn_t = (\underset{(.21)}{.382}B)\varepsilon ucgn_t + \gamma adn_t,$$

where γadn_t appears to be white noise. The within-sample variance of γadn_t is only 4 per cent less than that of εadn_t, as a comparison of rows 1 and 2 of Table II indicates. On the other hand, the form of model (AC.1) is economically plausible. Moreover, (AC.1) forecasts well in the post-sample period, yielding a 17 per cent improvement over the performance of (A.1).

We are now in a position to perform step (iv) of the procedure outlined in Section 3, the construction of models relating the original series. The evidence so far suggests that a unidirectional bivariate model is appropriate, with $UCGN_t$ causing ADN_t, but not the reverse. Substituting for εadn_t and $\varepsilon ucgn_t$ in (AC.1) from (A.1) and (C.1), appropriate forms for the final forecasting model can be identified. Estimation and deletion of insignificant higher-order terms yields the following bivariate model:

$$\text{(AC.2)} \quad (1+\underset{(.13)}{.327}B - \underset{(.16)}{.625}B^2)(1-B)ADN_t$$
$$= \underset{(.0025)}{.00665} + (\underset{(.21)}{.636}B + \underset{(.19)}{.317}B^5)UCGN_t + (1-\underset{(.19)}{.686}B^2)\eta adn_t,$$

$$\text{(C.2)} \quad (1-B)(1-B^4)UCGN_t = \underset{(.00055)}{.00126} + (1-\underset{(.12)}{.223}B^2 - \underset{(.13)}{.659}B^4)\varepsilon ucgn'_t.$$

Note that (C.2) is not identical to the univariate model (C.1) presented earlier. This is because (C.1) was estimated using a standard univariate Box-Jenkins program that used backforecasting to produce unconditional estimates, whereas all bivariate models had to be estimated with a more general (but less convenient) nonlinear least squares program that produces conditional, single-equation estimates [**2**, Sect. 7.1].[20] For most models, these procedures yield virtually identical estimates. Rows 4 and 6 in Table II indicate that (C.1) is slightly better than (C.2) within the sample, but it produces slightly worse forecasts in the post-sample

[20] See footnote 11.

period. Model (C.2) thus appears to be the appropriate one to use for post-sample comparisons.

The auto-correlograms of the residual series ηadn_t and $\varepsilon ucgn_t'$ are given in rows 2 and 3 of Table I. Both pass the standard single-series tests for whiteness. The cross-correlogram between these two series is given as row 4 in Table I. Several of the correlations for negative k suggest that further lagged values of $UCGN_t$ should be added to the right-hand side of (AC.2). A variety of experiments of this sort were performed in the course of identifying the model, however, and no significant or suggestive results were obtained. An examination of the correlations for positive k in row 4 of Table I shows that none exceeds one asymptotic standard error, $1/\sqrt{n} = .14$. The correlation at $k = +1$ is nonnegligible, however, and its size and location are suggestive. If the large contemporaneous correlation between the residual series is partly due to advertising causing consumption, one would expect the previous quarter's advertising to have some effect on current consumption. This effect should show up as a nonzero correlation between $\varepsilon ucgn_t'$ and ηadn_{t-1}. On the other hand, it is hard to rationalize taking the isolated nonnegligible correlation at $k = +4$ seriously. Thus the marginal term at $k = +1$ led us to identify and estimate the following model as a check on the (AC.2)/(C.2) structure:

$$\text{(CA.2)} \quad (1-B)(1-B^4)UCGN_t$$
$$= \underset{(.00090)}{.001885} - \underset{(.076)}{.121}(1-B)ADN_{t-1} + (1 - \underset{(.15)}{.162}B^2 - \underset{(.11)}{.684}B^4)\eta ucgn_t.$$

The series $\eta ucgn_t$ passes the standard tests. A comparison of rows 6 and 7 in Table II indicates that (CA.2) performs slightly better than (C.2) in both sample and post-sample periods.

We now turn to step (v) of our procedure, the evaluation of the post-sample forecasting performance of models fitted to the original series. Let us first consider models (C.2) and (CA.2). Use of the formal comparison test presented in Section 3 is ruled out here because, while the bivariate model, (CA.2), had a smaller forecast error variance at the 18 per cent level of significance, its mean forecast error was larger at the .1 per cent level. (These significance levels are based on one-tailed t tests on regression equation (4) in Section 3.) The overall post-sample mean-squared error for the bivariate model is only 5.1 per cent lower than for the univariate model, and neither of these tests suggests that this difference is significant at any reasonable level. We conclude, therefore, that the bivariate model, (CA.2), is not an improvement on the univariate model for aggregate consumption, (C.2); past advertising does not seem to be helpful in forecasting consumption.[21] We must accordingly retain the null hypothesis that aggregate advertising *does not* cause aggregate consumption.

[21] In earlier versions of this paper, we argued that this conclusion was strengthened because the negative coefficient of $(1-B)ADN_{t-1}$ in (CA.2) made no economic sense. Chris Sims has pointed out to us, however, that a negative coefficient is not all that implausible. Suppose that the main effect of aggregate advertising is to increase current spending on durables at the expense of future spending. Then, all else equal, a "high" value of past advertising would lead one to expect a "low" value of current consumption spending.

In contrast, Table II indicates that our bivariate model for aggregate advertising, (AC.2), forecast noticeably better than the univariate model, (A.1), reducing the post-sample MSE by some 26 per cent.[22] The post-sample forecast error series from both models had positive sample means. The Durbin-Watson statistic for equation (4), in Section 3, was 2.35 (20 observations), so no autocorrelation correction was indicated. Both coefficient estimates were positive, and the F statistic (with 2 and 18 degrees of freedom) corresponding to the null hypothesis that both population values are zero was 1.86, significant at the 18.4 per cent level.[23] In light of the discussion in Section 3, this means that we can reject the null hypothesis that the two models have equal mean-squared errors in favor of the superiority of the bivariate model at something less than the 9.2 per cent level of significance. This is hardly overwhelming evidence, but it does suggest that aggregate consumption is useful in forecasting aggregate advertising, and this indicates that consumption *does* cause advertising.

6. CONCLUSIONS

Applying the definition of causality discussed in Section 3, the analysis of Section 5 provides evidence that fluctuations in aggregate consumption cause fluctuations in aggregate advertising. No significant statistics suggesting that advertising changes affect consumption were encountered. Our empirical results are thus consistent with a model in which causation runs only from consumption to advertising.

Of course, any set of empirical results is in principle consistent with an infinite number of alternative models. In order to establish the value of the evidence we have presented, it is necessary to consider whether our results could have arisen from plausible alternative models with different causal structures.

As we noted in Section 5, our results are consistent with "instantaneous" causation from advertising to consumption.[24] All cross-correlograms between pairs of prewhitened series show high contemporaneous correlations. This suggests the possibility of an instantaneous or very short-term (within one quarter) relationship between advertising and consumption. But there is no way to tell if this relationship involves consumption causing advertising, advertising causing consumption, or a feedback structure involving both directions of causation. Thus, sudden unexpected changes in aggregate advertising may affect consumption within a quarter, but the finding that past advertising does not help in forecasting consumption indicates that such effects, if they exist, do not persist

[22] It is worth noting that the model built on the original variables, (AC.2), out-performs the model built on the prewhitened series (AC.1). This is consistent with specification error in the latter, as discussed toward the end of Section 3.

[23] From M. Ambramowitz and I. Stegun, *Handbook of Mathematical Functions* (Dover, 1972), equation (26.6.4), the significance level of an F-statistic with 2 and n degrees of freedom is given exactly by $[n/(n+2F)]^{n/2}$.

[24] It should be clear that the difficulty of interpreting contemporaneous correlations in causal terms is not particular to our approach to testing for causality or to our data set.

over time intervals that are substantial relative to a calendar quarter. It seems implausible to us that advertising affects consumption in this fashion.

As Sims [**20**] has pointed out, if one variable, X_t, is used to stabilize another, Y_t, optimally over time, the resultant time series can show spurious causation from Y_t to X_t. But this does not seem likely to be a problem here. It is somewhat implausible to think that uncoordinated advertising decisions lead the business sector to act "as if" accurately stabilizing aggregate consumption. But more importantly, if the structural effect of advertising on consumption were positive, and if the exogenous disturbances to consumption were positively serially correlated, the optimal control hypothesis would imply negative, not positive coefficients on lagged consumption in model (AC.2).

Though our data set was superior to those previously employed to study the aggregate advertising/consumption relation, it was not entirely satisfactory. First, it would have been preferable to have worked with advertising data that had not been seasonally adjusted. On the other hand, as pointed out in footnote 19, seasonality problems here should have biased our estimates toward finding causation from advertising to consumption. Second, it is at least plausible that *ADN* is more infected with measurement error than *UCGN*. As Sims [**20**] has shown, this can lead to a spurious causal ordering in the direction we find. However, it seems unlikely to us that measurement error in ADN_t is sufficiently large relative to its quarter-to-quarter variation to have significantly affected the results reported here.

Finally, the total sample of 80 observations was not as large as would have been desirable. Given the importance of post-sample testing in our approach, a post-sample period of more than 20 observations might have permitted more precise inferences. Were we to do this study again, we would probably divide the data more evenly between sample and post-sample periods for this reason. Of course, this problem relates to the strength of our conclusions, not directly to the pattern of causation we detect.[25]

In short, causality testing with typical economic data remains at the frontier of econometric work and is hence a rather non-routine affair. Nevertheless, we believe that the results discussed above showing that fluctuations in past aggregate consumption appear to influence aggregate advertising, but not vice-versa, are valid at the significance level quoted.

Moreover, our experience with the test for causality proposed in Section 3 has left us confident of its utility. Its first desirable feature is the focus on the original variables rather than the pre-whitened (residual) series. In the application in Section 5, steps (iv) and (v) yielded much stronger evidence than did the analysis of pre-whitened series in steps (ii) and (iii). The second desirable feature of our approach is its stress on out-of-sample forecasting performance. We discussed the complexities involved in optimal use of out-of-sample data in Section 3. Sample

[25] In addition to these problems, we cannot rule out the possibility that our results were generated by a structure in which advertising and consumption both depend on some omitted third variable. But Sims [**20**] has shown that conditions under which spurious causal orderings can arise in this fashion are rather implausible.

data mining (leading to specification error) and structural instability can lead to difficulty in obtaining useful causal inferences with the methodology proposed here. However, we find this possibility distinctly preferable to the spurious inferences that these problems can easily produce when out-of-sample verification is not employed. Similarly, restricting causal hypothesis testing to a separate out-of-sample period clearly decreases the number of degrees of freedom available for such testing; on the other hand, only then can one be really sure that none of those degrees of freedom have been "used up" in the model identification and estimation process.

University of Texas, Austin,
University of California, San Diego,
and
Massachusetts Institute of Technology

Manuscript received June, 1977; revision received July, 1979.

APPENDIX

The CBS advertising spending estimates are used here instead of the PII for two reasons. First, changes in media coverage in the PII cause a break in 1971.[26] Second, within the 1953–1970 period, the media covered by the PII become increasingly unrepresentative over time.[27]

In [**15**, App. B], CBS estimates of quarterly movements of national advertising spending in newspapers, magazines, business papers, outdoor media, network television, spot television, network radio, and spot radio were employed to extend Blank's [**1**] series through 1967.[28] For this study, we obtained more recent CBS estimates of quarterly spending in all these media except business papers and outdoor media for the 1966–1975 subperiod,[29] along with current McCann-Erickson estimates of annual spending totals in these media for the entire

[26] See the May and June, 1971 issues of the *Survey of Current Business.* A similar break occurred between 1952 and 1953 [**23**, p. 8].

[27] PII covered network radio and television but did not cover the spot markets in these media. (Spot television was added in 1971.) By 1966, national advertising spending for spot television was two-thirds that for network television, while spending in spot radio was more than four times that for network radio [**15**, p. 8].

[28] National advertising is prepared centrally and disseminated to several localities, while local advertising is prepared and disseminated in the same locality. Local advertising is largely done by retailers, while national manufacturers are the dominant national advertisers.

[29] Spending in business papers was excluded because we did not expect it to be causally related to household consumption spending. Outdoor media had to be dropped because CBS had stopped preparing quarterly estimates. The CBS series were seasonally adjusted at the source using (basically) the Census X-11 program. The sources used by CBS in preparing the earlier data are discussed in [**1**; **15**, App. B]. The more recent estimates of quarterly movements are based on information from the Television Bureau of Advertising, Broadcast Advertisers Reports, *Television/Radio Age*, the Radio Advertising Bureau, the Newspaper Advertising Bureau, Publishers' Information Bureau, and a cooperative service commissioned by the major radio networks.

1956–1975 period.[30] The quarterly totals reported in [**15**, App. B] were used for the 1956–1965 subperiod. The quarterly flows for each medium were re-scaled, where necessary, so that annual averages equaled the McCann–Erickson annual totals. The six resultant series were used, along with quarterly population from various issues of the *Survey of Current Business*, to obtain *ADN*.

A set of annual cost-per-million (*CPM*) indices, which reflect changes in both media costs and audience sizes, were obtained from McCann-Erickson for the media covered by *ADN* for the 1960–1975 subperiod. These were linked to the *Printer's Ink* indices reported in [**15**, App. A] at 1960. The six *CPM* indices were then interpolated, using a linear method that ensured that the averages of the quarterly indices equaled the annual value.[31] The six current dollar spending series were deflated by the resultant quarterly *CPM* indices, and the deflated totals were used, along with the population series, to obtain *ADR*.

REFERENCES

[**1**] BLANK, D. M.: "Cyclical Behavior of National Advertising," *Journal of Business*, 35 (1962), 14–27.

[**2**] BOX, G. E. P., AND G. M. JENKINS: *Time Series Analysis.* San Francisco: Holden-Day, 1970.

[**3**] BOX, G. E. P., AND D. A. PIERCE: "Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models," *Journal of the American Statistical Association*, 65 (1970), 1509–1526.

[**4**] CLARKE, D. G.: "Econometric Measurement of the Duration of the Advertising Effect on Sales," *Journal of Marketing Research*, 13 (1976), 345–357.

[**5**] DAVIES, N., C. M. TRIGGS, AND P. NEWBOLD: "Significance Levels of the Box-Pierce Portmanteau Statistic in Finite Samples," *Biometrica*, 64 (1977), 517–522.

[**6**] DURBIN, J.: "Testing for Serial Correlation in Least-Squares Regression When Some of the Regressors are Lagged Dependent Variables," *Econometrica*, 38 (1970), 410–421.

[**7**] EKELUND, R. G., AND W. P. GRAMM: "A Reconsideration of Advertising Expenditures, Aggregate Demand, and Economic Stabilization," *Quarterly Review of Economics and Business*, 9 (1969), 71–77.

[**8**] GRANGER, C. W. J.: "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods," *Econometrica*, 37 (1969), 424–38.

[**9**] GRANGER, C. W. J., AND P. NEWBOLD: *Forecasting Economic Time Series.* New York: Academic Press, 1977.

[**10**] HOUTHAKKER, H. S. AND L. D. TAYLOR: *Consumer Demand in the United States*, 2nd Ed. Cambridge: Harvard University Press, 1970.

[**11**] KOTLER, P.: *Marketing Management*, 3rd Ed. Englewood Cliffs: Prentice-Hall, 1976.

[**12**] MADDALA, G. S., AND A. S. RAO: "Tests for Serial Correlation in Regression Models with Lagged Dependent Variables and Serially Correlated Errors," *Econometrica*, 41 (1973), 761–764.

[30] The McCann-Erickson totals include both media charges and production costs. These estimates appear at intervals in *Advertising Age*. See also [**22**, Ch. I] and recent numbers of the *Statistical Abstract of the United States*.

[31] Let $X(t)$ be the value of some series in year t, and let $x(i, t)$ be the interpolated value for quarter i of that year. In the interpolation method employed, $x(1, t)$ and $x(2, t)$ were found by linear interpolation between $X(t-1)$ and an adjusted number $X'(t)$, and $x(3, t)$ and $x(4, t)$ were similarly based on $X'(t)$ and $X(t+1)$. $X'(t)$ was selected for each year for each series so that the average of the $x(i, t)$ equalled $X(t)$. This makes all the $x(i, t)$ linear functions of $X(t-1)$, $X(t)$, and $X(t+1)$. (Ordinary linear interpolation was used for 1975.) Using standard tests for homogeneity of means and variances of percentage changes ending in each of the four quarters, this method performed well relative to a variety of alternative average-preserving interpolation techniques.

[13] 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*, 72 (1977), 11–22.

[14] PIERCE, D. A., AND L. D. HAUGH: "Causality in Temporal Systems: Characterizations and a Survey," *Journal of Econometrics*, 5 (1977), 265–293.

[15] SCHMALENSEE, R.: *The Economics of Advertising*. Amsterdam: North-Holland, 1972.

[16] SIMON, J. L.: *Issues in the Economics of Advertising*. Urbana: University of Illinois Press, 1970.

[17] SIMS, C. A.: "Money, Income, and Causality," *American Economic Review*, 62 (1972), 540–552.

[18] ———: "Seasonality in Regression," *Journal of the American Statistical Association*, 69 (1974), 618–626.

[19] ———: "Comment," *Journal of the American Statistical Association*, 72 (1977), 23–24.

[20] ———: "Exogeneity and Causal Ordering in Macroeconomic Models," in *New Methods in Business Cycle Research*, ed. by C. A. Sims. Minneapolis: Federal Reserve Bank of Minneapolis, 1977.

[21] TAYLOR, L. D., AND D. WEISERBS: "Advertising and the Aggregate Consumption Function," *American Economic Review*, 62 (1972), 642–655.

[22] U. S. BUREAU OF THE CENSUS: *Historical Statistics of the United States: Colonial Times to 1970*. Washington: U.S. Government Printing Office, 1975.

[23] VERDON, W. A., C. R. MCCONNELL, AND T. W. ROESLER: "Advertising Expenditures as an Economic Stabilizer, 1954–64," *Quarterly Review of Economics and Business*, 8 (1968), 7–18.

[24] WALLIS, K. F.: "Seasonal Adjustment and Relations Between Variables," *Journal of the American Statistical Association*, 69 (1974), 18–31.