April 1976

#85

HAVE GEOMETRIC LAG HYPOTHESES OUTLIVED THEIR TIME? SOME EVIDENCE IN A MONTE CARLO FRAMEWORK

by

John F. Wilson

Most distributed lag models have almost no or only a very weak theoretical underpinning. Usually the form of the lag is assumed a priori rather than derived as an implication of a particular behavioral hypothesis. Exceptions to this statement are the adaptive expectations and partial adjustment models, which in part explains their recent popularity. But even here, the theoretical rationalizations offered are often only skin deep.

Zvi Griliches -
"Distributed Lags: A Survey"

"Have Geometric Lag Hypotheses Outlived Their Time? Some Evidence in a Monte Carlo Framework"

John F. Wilson*

I. Introduction

It is probably fair to say that there is today universal agreement that economic variables do not adjust instantaneously to their determinants. Full adjustment is in some sense distributed over time. A second proposition which would probably meet little opposition is that discerning the correct shape and length of an adjustment process is a very ticklish matter. Although the recent history of distributed lag estimation has witnessed great progress since the initial work undertaken by Fisher and Tinbergen in the 1930's, a great deal remains to be done.

The intention of the present paper is to illustrate, in a Monte Carlo framework, the results of applying several of the currently fashionable distributed lag estimating techniques to a body of real world data on which various known lag distributions have been imposed. In particular, the argument will be made (using the geometric lag hypothesis as an example) that investigators who begin with easily estimatable, but unduly rigid, hypotheses run the risk of obtaining misleading results. The conclusion which emerges is that certain types of frequently-used initial hypotheses have to some degree been "obsoleted" by the development in recent years of more flexible techniques for evaluating distributed lags.

---

*Economist, International Finance Division, Board of Governors of the Federal Reserve System. The views expressed herein are solely those of the author and do not necessarily represent the views of the Federal Reserve System. The author is grateful to Scott Brown and Irene Cavanagh for capable research assistance in preparing the statistical material included in this paper.

II. Basic Lag-Estimating Techniques: A Thumbnail Sketch

Prior to the major contribution in 1954 by Koyck [12], an investigator studying distributed lag processes had only a narrow range of feasible options. Among these, simple OLS procedures may have been the least distasteful, despite the problems raised by time-series collinearity. Koyck's discovery that, in a certain type of bivariate equation, a simple transformation could replace an infinite string of lagged regressors with a single lagged dependent variable revolutionized the field.[1/] Although it was necessary for Koyck to assume that the coefficients on the lagged regressor terms declined exponentially, estimation of geometrically declining lag coefficents subsequently enjoyed (and still does) a great vogue. In the absence of more flexible procedures, there was initially hardly any gain to arguing that the world might not be governed entirely by exponential processes. Not much time had to pass, however, before Cagan [3] and Nerlove [16] came up with two rationales which lent more economic credence to Koyck's mathematical expedient.[2/] This lag technique is still very much with us.

1/ Koyck himself made only modest claims, writing [12, p. 4]: "This study, it is hoped, is one minor step in a great number of successive steps still ahead in the field of estimating structural economic relations from time-series data."

2/ Cagan's "adaptive expectations" hypothesis concerns the adjustment rate of the regressor in such a relationship. Nerlove's "partial adjustment" mechanism hypothesizes a similar adjustment process for the regressand. In their simple forms, both models result in Koyck-type estimating equations, which leads to ambiguity about which hypothesis is being tested. The adaptive expectations approach also introduces serial correlation into the error term, even if it were absent in the structural hypothesis. The partial adjustment approach does not. Other complications are noted in Section VI below.

In his 1967 survey on distributed lags, Griliches [8, pp. 24-25] summarized the case for geometric lags as follows: "Its main advantage is ease of estimation -- everything depends on only one additional parameter. This is done, however, at the cost of forcing a particular form of the lag on the data."

The second widely used distributed lag technique is that developed in 1965 by Almon [1], and its working assumption is that a lag structure lies along some low order polynomial. Following the appearance of Almon's paper, some researchers began to "think polynomial," and enjoyed the advantage of being able to develop separate distributions for as many variables as they chose to include in a given function.[3/] Polynomial interpolation is not an unmixed blessing, however, because use of the Almon technique implies a willingness on the researcher's part to sort through various (unknown) lag lengths and curve degrees. He must be ready to "search" more actively for the lag structure than before.

The method recently developed by Shiller [22] carries the evolution of estimating techniques to yet a higher degree of flexibility. Shiller's procedure was developed from Bayesian priors by assuming that a linear combination of the coefficients (in this case coefficient <u>differences</u> of some degree) in a lag distribution are normally distributed with a zero mean and some variance. If this variance is also zero, Shiller

---

3/ This is also not entirely impossible with geometric assumptions.

shows that the method is equivalent to the Almon procedure.[4/] However, the advantage of the method is that it is stochastic, which makes it possible for estimation results to deviate from the investigator's implicit expectations. That is, in the search for the "true" lag distribution, even if the estimator prior should misspecify the order or shape of the true curve, there is still a possibility that the regression results can approximately identify the correct pattern.

While Shiller's estimator was originally derived in a Bayesian framework, it can also be described in terms of Theil-Goldberger type mixed estimation, as has been done by Shiller [23], Wilson [25] and Maddala [14]. A recent note by Taylor [24] also shows the equivalence between the two. Judging from the attention paid to it in the recent literature, this procedure is rapidly gaining in popularity.

A fourth estimator explored in this paper is the ridge regression technique. Originally proposed in two articles by Hoerl and Kennard [9, 10] simply as a method for overcoming matrix collinearity, ridge methods are also finding application in distributed lag problems. Recent examples can be found in notes by Rappoport [19] and Maddala [14]. In the latter paper, the Shiller estimator and various others are treated as special cases of ridge methods.

4/ It will be recalled that for a polynomial of degree "d", the "d + 1"th differences of the terms equal zero. Shiller's priors on any given level of coefficient differences, therefore, implicitly specify a polynomial curve one degree lower than the difference degree. References to Shiller's method in this paper will generally be in terms of "polynomial equivalents."

In most ways the foregoing discussion does little more than scratch the surface of the literature on lags, but it is assumed that the reader is familiar with the derivations of these estimators and examples of their use. Beyond the four methods explored in this paper, a variety of other more or less feasible techniques is available. Solow, for instance, explored Pascal distributions as lag priors; Jorgenson developed a rational distributed lag model (of which Almon and geometric lags are special cases); spectral methods are associated with Hannan; and Shiller [21] has recently extended his work by developing priors on differences in the logarithms of lag coefficients for the case where all coefficients are expected to be positive.5/ Along the way, certain more exotic forms have also emerged. Schmidt [20], for instance, proposed an estimator which is a sort of hybrid of the Almon and geometric methods. Corradi and Gambetta [4] advocate the use of spline functions. In this welter of developments, certain facets of distributed-lag estimation have become rather complicated. Just how complicated they can become will be appreciated by anyone who has read the recent book by Dhrymes [5] on the subject.6/

Each of the above-noted methods of estimating distributed lags imposes certain prior information on a body of data. This information implies something about the character of the lag curve which the

---

5/ In his recent paper [21, p. 12] Shiller also shows the equivalence of his new prior, applied in a certain way, to the geometric estimator. The advantage of this new version is that it could be used to estimate a (geometric) lag-pattern on a single variable.

6/ Dhrymes notes a variety of reasons why geometric lag hypotheses have gained "wide currency," but says also that one "should not turn to a theoretically deficient model simply because the estimation problems it presents can be easily tackled." [5, p. 55] investigator expects to find.[7/] But it is also clear that the evolution over time of more flexible distributed-lag techniques has been characterized by the conjuring of increasingly less restrictive prior restrictions concerning the structure of the lag process. The Almon method, for instance, rests on a less rigid view than the Koyck about what a true lag structure might look like. The prior information imposed by the Shiller technique is, in turn, less restrictive than the Almon priors. Each of these three methods, nonetheless, exploits priors which contain information bearing on the relationship between "adjacent" coefficients in the lag distribution. Ridge regression brings to bear information which is somewhat different in this regard. The prior imposed by this technique has little to do with the relation between a lag coefficient b_i with an adjacent coefficient b_{i+1}, but rather with the allowable magnitude of the product of the entire coefficient vector, b'b.[8/] Correspondingly, the ridge technique is better viewed as an expedient to break collinearity than as a method derived on the basis of an articulated hypothesis about the form of some particular lag distribution.

7/ Each method "works" statistically, however, largely due to data transformations which disrupt the massive collinearity which makes OLS estimation so difficult with many time series.

8/ The ridge priors are introduced by adding information, in the form of μI, to the moment matrix (or submatrix of lagged regressors) of the data prior to inversion. The regression thus takes the form: $b=(X'X+\mu I)^{-1}X'y$, and b'b is suppressed toward zero as the value of μ is increased. For this reason, when ridge methods are used and μ is set at a high level, individual coefficients are also "urged" toward zero. The effect is in some ways analagous to far-endpoint priors under the Almon or Shiller methods.

III. Estimator Evolution and the Persistence of Older Forms: Some Examples

The purpose of the following sections is not to examine the virtues of these estimators *per se*. But if it is true, as the author believes, that the recent history of estimation techniques represents a true evolution, such an evolution implies that older methods are to some extent superseded by new ones. Behind this argument lies the observation that newer and more general estimators often at least approximately encompass the older ones as special cases. For example, although a geometric lag shape is not a simple low-order polynomial, it is encompassed by the general category of rational distributed lags. A polynomial curve of any given degree is also just a special case of all higher order curves in which, as it happens, certain coefficients are equal to zero. This fact suggests that the proper way to identify a (polynomial) lag of given degree -- leaving length questions aside for the moment -- might be to use programming which "assumes" the curve is in fact of higher degree than suspected. Such a deliberate high-side misspecification can in theory and fact identify lower-order lag structures.[9/] Because the Shiller estimator can be roughly described as a "stochastic Almon," the same comments are applicable to the use of this technique.

A general argument can thus be made that in order to demonstrate the correctness of a hypothesis which assumes a particular lag form, corroboratory evidence from a more general hypothesis (and its estimators) should be sought. Most frequently, however, an investigator will postulate

---

9/ Monte Carlo results obtained by the author bear this out. Known 2nd degree curves of various sorts (and with various error variances) were successfully identified using not only 2nd, but also 3rd and 4th degree assumptions when the lag length is also correctly found.

a model, derive the corresponding lag estimator and make some runs in which "good" results are obtained. By implication, said "good" results lend support to the initial hypothesis. This does not really follow, especially in the absence of a strong a priori case in favor of the hypothesis itself, without regard to the estimation results. For instance, an investigator who has settled on a 2nd degree curve may remark that the equation fit was better than with other curves or that coefficient significance was acceptable, but often without telling the reader whether the results gotten in tests with a higher order curve happen to have even looked like a 2nd degree polynomial.[10/] However, if supplementary tests do not support such conclusions, it seems fair to ask that this negative evidence be taken into account.[11/]

As a prime case in point, Koyck-type structures have been popular with the profession for the last twenty years. Once a lagged dependent variable has been jiggled to the RHS of an equation, one estimates geometric lags. More precisely, one estimates only geomtric lags. Such a procedure makes rather uncharitable assumptions about the capacity of the data to tell its own story. Although one might have thought that newer techniques would have replaced this approach, this does not in fact seem to be the case. Just in the body of literature with which this

10/ In multivariate regressions, especially those on trended data, the difference in fit is usually marginal at best. The results shown in Tables 1-3 below illustrate this point.

11/ In practice, omitted variables and other functional misspecifications can produce such results, even when the low-order lag process exists. The point is, rather, that in the literature such evidence is usually omitted altogether.

writer is most familiar -- the area of international trade studies -- numerous Koyck-type models have been estimated in the years since the Almon method became generally available. Among these are studies by Branson [2], Kwack [13], Prachowny [17], Gregory [7], Rao [18], Hooper [11], Miller-Fratianni [15], and most recently Goldstein and Khan [6]. One suspects that, in the absence of *a priori* reasons why a geometric lag pattern should obtain, the main reason for the choice of this technique was ease of estimation. Sorting through an abundance of curves and lag lengths is, after all, rather messy and tiring compared with the simplicity of the alternative.[12/]

From the above it should be clear that the writer does not find such a solution to be satisfying. The remainder of this paper will therefore be devoted to introducing evidence drawn from a recent Monte Carlo experiment to support the following two propositions:

a) Even if the true lag structure is *not* in fact characterized by geometric decay, the results of Koyck-type estimation can (spuriously) tend to support the Koyck-hypothesis.

b) Supposing, in contrast, that the true structure *is* characterized by a geometric lag structure on one or more of the regressors, other forms of lag estimation are capable of approximately identifying such patterns, and there is thus no need to *begin* with a Koyck-type hypothesis. Experiments with such a structure have been made using Almon, Shiller and ridge estimators.

12/ The paper by Miller-Fratianni is especially interesting in that *both* Almon and geometric-type estimations are made. The results are strikingly different, but the authors offer no comment on why this might be or what it suggests for the validity of either set of estimates.

IV. Equation Forms and Error Structure

Drawing on a data-bank assembled by the author in connection with a disaggregated study of U.S. import demand [25], three variants of the following equation were constructed:

$$DV_t = k + a \ln (Y)_t + \sum_{i=o}^{n} b_i \ln \left(\frac{P_f}{P_d}\right)_{t-i} + e_t$$

This is a simplified replica of a fairly typical U.S. import demand relation, and the right-hand variables represent actual historical data for the period 1958.I-1971.IV.[13] Variable definitions follow below:

- DV = dependent variable (sum of RHS components, including the random error term);
- Y = United States GNP in billions of constant dollars (nominal GNP divided by the implicit deflator);
- P_f = Implicit deflator for imports of goods and services;
- P_d = Deflator for U.S. private, non-farm output;
- e_t = random error term.

Since this experiment is for illustrative purposes only, no lag structure was assigned to the income term.[14] No error component was added to right-hand variables, and the error term e_t does not follow an autoregressive scheme. Using results from typical aggregate U.S. import equations, the "known" parameters (elasticities) in the above equation were set as follows:

$$k = -5.0$$
$$a = 1.2$$
$$\Sigma b_i = -2.0$$

13/ The net estimating "sample," allowing for various lag constructions, encompasses the 1962.III to 1971.IV period. Price terms were set on a 1963 base, although any other base could be used equally well.

14/ Most empirical trade studies not using Koyck transformations have in any case "found" lag structures to be much shorter on income than on price terms in such an equation.

The basic approach thus is to assume that there is reasonably reliable information on certain aspects of such an equation (e.g., long-run elasticities), but little on the shape or time-span of adjustment. The investigator searching for such information might therefore try a variety of lag techniques, most of which will, by definition, be incorrectly specified. On the basis of the regression results he will have to conclude what he can about the unknown process which generated the data.

To illustrate some cases which might occur, three different forms of lag distributions were assigned to the relative price term of the basic equation. These are described below:

Case 1: No lag structure. The parameter values corresponding to this case are $b_o = -2.0$, and $b_i = 0$ for all $i>0$.

Case 2: A rudimentary, truncated lag structure which is not geometric in type. Here, for the basic case, $b_o = -1.5$; $b_1 = -0.5$; and $b_i = 0$ for all $i>1$. An alternative case will also be discussed.

Case 3: Geometric decay in the lag structure on the relative price term. The coefficient vector was set up to obey the relation

$$b_i = \lambda^i b_o, \text{ where } b_o = -0.8 \text{ and } \lambda = 0.6.$$

This process yields the following values for the individual lag coefficients:

$b_0 = -.800$	$b_6 = -.03733$
$b_1 = -.480$	$b_7 = -.022395$
$b_2 = -.288$	$b_8 = -.01343$
$b_3 = -.1728$	$b_9 = -.00806$
$b_4 = -.10368$	
$b_5 = -.0623$	$\Sigma b_i = -1.987995$

The error term e_t in each equation for cases 1-3 was generated by the standard Fortran Subroutine Gauss, which produces a normally distributed variate, with mean and standard deviation specified by the user. Therefore $e_t \sim N(\mu, s^2)$, where for the results tabulated $\mu = 0$ and $s_e = .020$.15/

One further comment should be added before the empirical results are described. In a full scale Monte Carlo experiment a large number of replications are needed to approximate the large-sample properties of the estimator. Due to the limited scope of this study -- which, however, still required the tabulation of a goodly number of equations -- the results summarized in the following tables are based on only ten runs with the data. (The correlation matrix of current and lagged variables included in the regressions is shown in the Appendix.)

15/ Increasing the level of s creates rising levels of "noise" in the dependent variable, which has detrimental effects on the fitting properties of the equation. A range of values from s = .002 to s = .040 was tried. The value .020 was chosen for tabulation largely because the resulting equation fits were about the same as those produced by aggregate import equations.

V. Estimation Results

Case 1: No lag distribution.

Letting Y be the log of real income and P stand for the log of relative prices, the equation to be estimated has the form:

$$DV_t = -5.0 + 1.2Y_t - 2.0P_t + e_t \quad (s_e = .020)$$

Basic estimation results for this equation are shown in Table 1 and Graph 1 on the following pages.16/ From these it can be seen that a simple, correctly specified equation (i.e., no lags) estimates this structure quite well. (In fact, in terms of fit, long-run price elasticity and estimate of the constant, all of the equations are satisfactory.) The other two unconstrained OLS estimates each come up with results close to the true values of a, k, and b_o, although the 9 quarter misspecification shows the highly erratic coefficient pattern typical of such regressions.

If a geometric lag is incorrectly hypothesized and a lagged dependent variable is added to the RHS of such an equation, the results are still satisfactory in terms of the price structure. The mean coefficient on the lagged dependent variable comes out very low (.0340), and in all runs this coefficient was found to be insignificant. This suggests that *if* the true structure is of the restrictive Case 1 type, but the researcher incorrectly assumes a geometric type lag distribution exists, the error can be caught by the estimation results.17/

16/ Estimates shown in all tables are the means of the results from the 10 runs. To conserve space no measure of coefficient significance is shown, but comments on this point will be made in the text.

17/ That is, so long as there is no autocorrelation in the error structure. Griliches /8, pp. 33-34/ cites results from an equation similar to that in Case 1. His equation is $y_t = ax_t + u_t$, but u_t follows a first-order autoregressive scheme. The (erroneous) introduction of y_{t-1} as a regressor in this structure produced a significant coefficient. Griliches therefore concludes that "the partial adjustment model will work even if it is wrong."

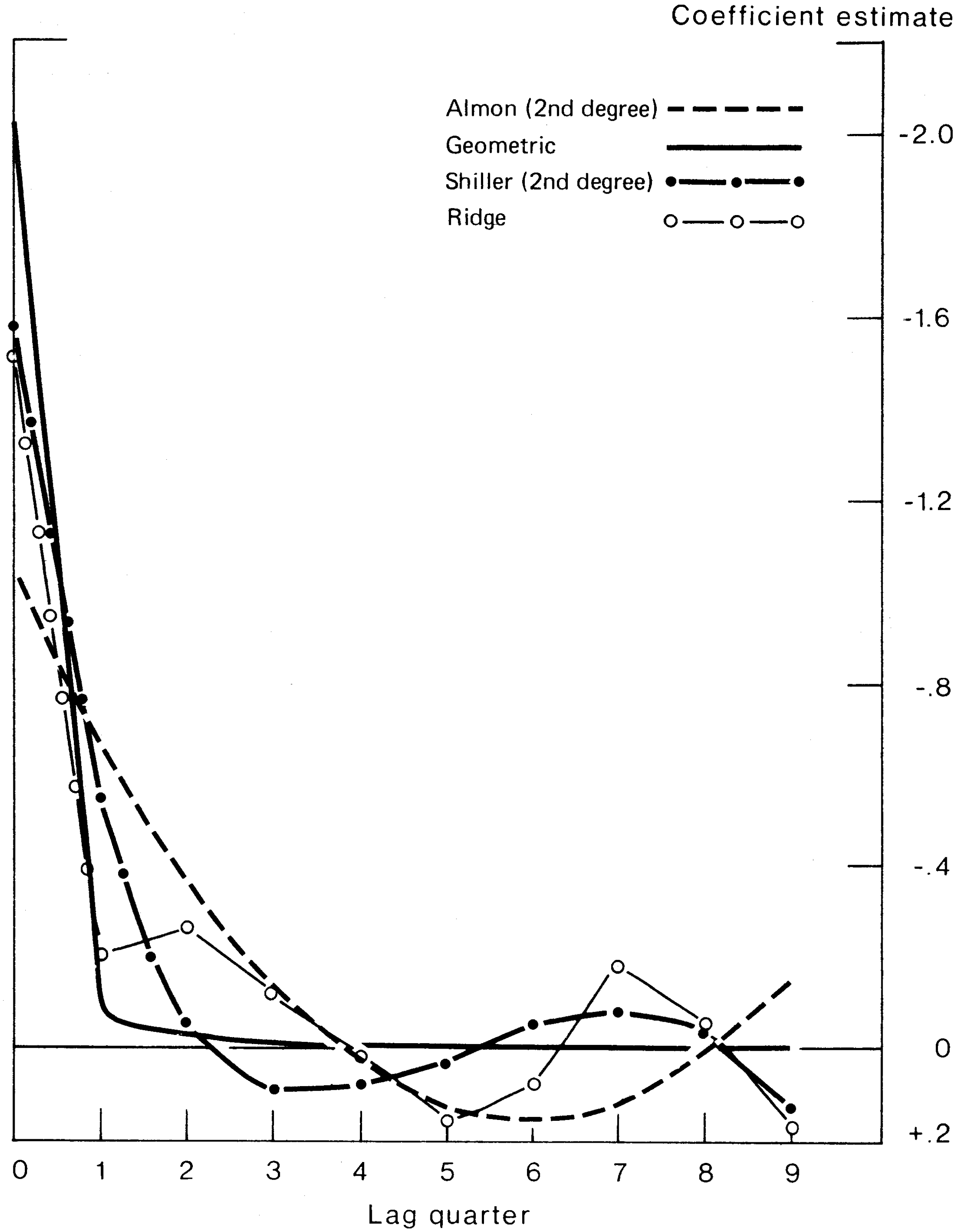
Table 1. Estimation Results for Sample Equation 1: No Lag Structure

True Equation and Estimation Type	a	b_0	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	c3/	k	$\bar{R}^2$	Σb_i
True Equation	1.20	-2.00	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0		-5.0		-2.00
OLS															
No Lag	1.1922	-2.0693											-4.9505	.9920	-2.0693
1 Q Lag	1.1902	-2.0246	-.0529										-4.9376	.9920	-2.0775
9 Q Lag	1.1917	-2.0409	.2987	-.3515	.1155	-.2472	.4115	-.0329	-.0712	-.2723	.2440		-4.9472	.9918	-1.9463
KOYCK (implied)	1.1466	-2.0215	(-.0687)	(-.0023)	(-.0001)	(-0.0)	(-0.0)	(-0.0)	(-0.0)	(-0.0)	(-0.0)	.0340	-4.7496	.9921	-2.0927
(long-run)	(1.1869)	(-2.0927)													
ALMON - 9 Q															
2nd deg. curve	1.1990	-1.0472	-.6759	-.3724	-.1376	.0237	.1283	.1591	.1218	.0099	-.1571		-4.9934	.9903	-1.9474
3rd deg. curve	1.1989	-1.4219	-.6235	-.1062	.1215	.1453	.0510	-.0757	-.1488	-.0828	.2083		4.9944	.9915	-1.9328
SHILLER1/ - 9 Q															
1st deg. curve	1.1996	-1.4131	-.6208	-.1533	.0510	.1041	.0654	-.0257	-.0917	-.0308	.1107		-4.9999	.9916	-2.0042
2nd deg. curve	1.1942	-1.5814	-.5519	-.0556	.0895	.0809	.0246	-.0467	-.0898	-.0398	.1239		-4.9636	.9918	-2.0463
3rd deg. curve	1.1897	-1.7109	-.4466	-.1144	.0766	.0294	-.0094	-.0361	-.0537	-.0274	.0895		-4.9343	.9918	-2.2030
RIDGE2/ - 9 Q	1.2161	-1.5038	-.2005	-.2658	-.1203	.0190	.1619	.0791	-.1836	-.0591	.1728		-5.1057	.9893	-1.9003

1/ Expressed in polynominal curve "equivalents". k = .05
2/ μ = .001
3/ Lagged dependent variable included.

GRAPH 1

Coefficient patterns for equation with no lag structure



A comparison of the Almon and Shiller estimates of the Case 1 equation is also instructive. All five of the cases shown have been deliberately grossly misspecified.[18/] None of these equations errs much in estimating the income coefficient or the constant. However, there is, for each estimator considered separately, a clear improvement in the estimate of b_o as the curve prior increases. It can also be seen that for the coefficients b_1-b_9 (whose true value is zero) results generally improve for increases in this same prior. As between the two estimators, for any given explicit or implicit curve assumption the Shiller results are clearly superior for most coefficients.[19/] Further, a number of the Almon coefficients are significant, whereas beyond b_o the Shiller coefficients are almost uniformly insignificant at accepted levels.

Results obtained by the ridge regression have the same erratic, disappointing features as those obtained by the nine-quarter OLS function.[20/]

18/ In Tables 1-3 only the nine-quarter results are shown for Almon, Shiller and ridge specifications. In Table 3 these lag lengths are approximately correct; in the others they are misspecifications. The author has also obtained less extreme five and seven quarter misspecifications which show similar features. The value of "k" (representing the ratio of the standard deviation of the random error in the structure to that of the prior) has been set at .05 for the tabulated Shiller equations. This value is about half that which would result from the rule of thumb suggested by Shiller /22, p. 779/.

19/ The exceptions generally occur at points where the Almon curves cross the zero line.

20/ There is no well-defined theory about the proper setting for μ in ridge regression. Hoerl and Kennard suggest a search procedure. Maddala, after trying values of 0.1 and 0.005, concludes that the value "has to be really very low." (p.11) Values used by the author ranged from 1.0 down to .001, with results from the lowest setting shown in the tables. Higher settings tended to depress all coefficients in the distribution toward zero, irrespective of their true values.

The equation considered under Case 1 has some theoretical interest, but is the most implausible of the structures which might obtain in the real world and is of limited usefulness. Most researchers would expect to find some sort of lag structure in such an equation, even if they had no clear notion about the exact nature of the lags. Let us therefore see what happens when even a simple lag-structure on prices is tacked onto the equation.

Case 2: Rudimentary lag on prices.

The specimen equation for this case is written as follows:

$$DV_t = -5.0 + 1.2Y_t - 1.5P_t - 0.5P_{t-1} + e_t \quad (s_e = .020)$$

There is again no need to persuade anyone that such a structure is plausible, but it is interesting to note the difference in estimation results when even this one period of lagged influence is assumed. Table 2 and Graph 2 on the following pages summarize these findings.

As in the no-lag case, all equations do reasonably well in identifying the long-run price response. The correctly specified equation (OLS-1Q price lag), as expected, also performs well, and the nine-quarter OLS results have the same erratic properties as seen earlier. Since the first OLS equation contains no lagged terms, it has no choice but to "assign" the full price response to the current quarter.

More interesting results are obtained with the (again, incorrect) geometric specification. The estimate of b_0 is overstated and that of b_1 is understated; succeeding coefficients indeed go to near zero.

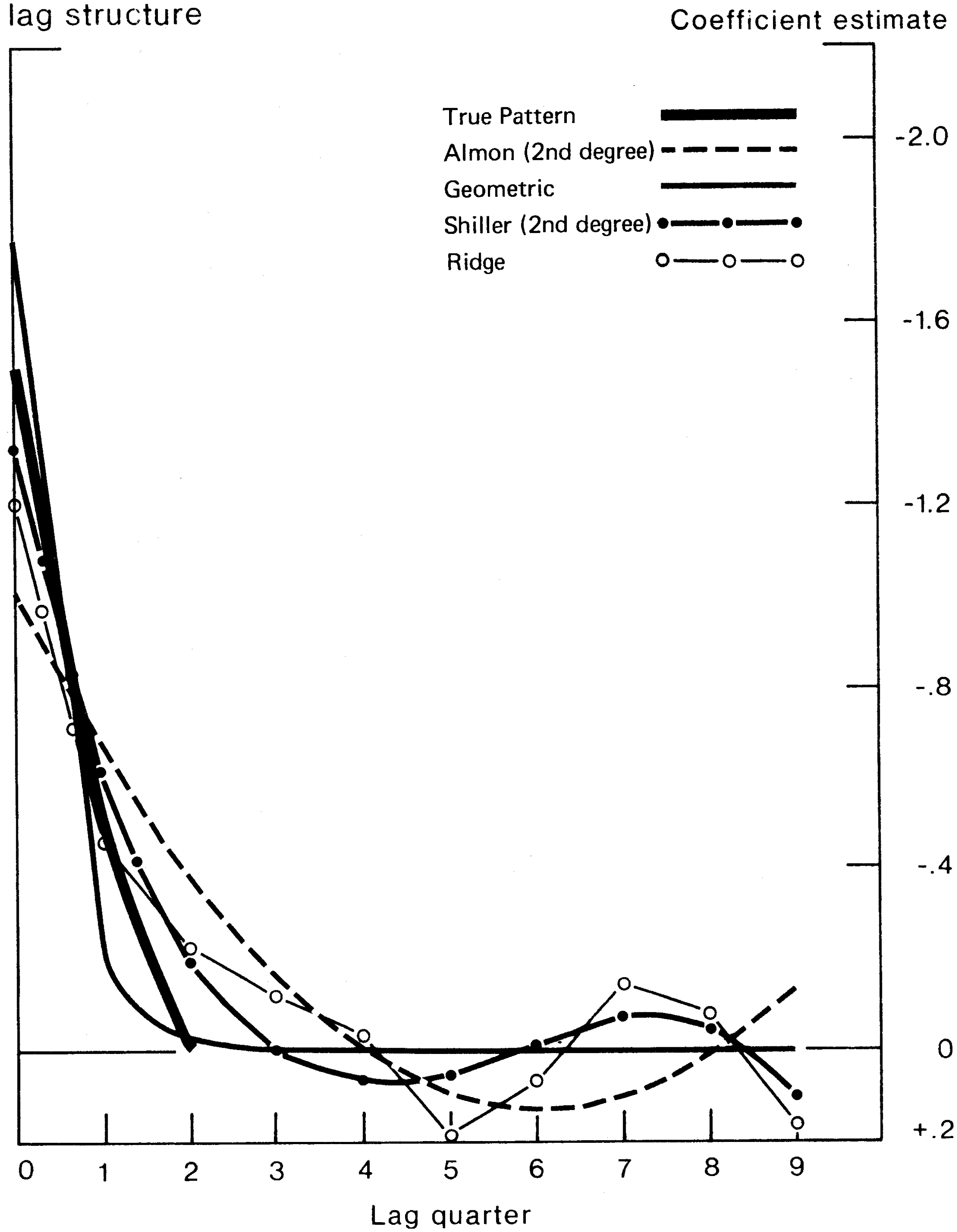
Table 2. Estimation Results for Sample Equation 2: Rudimentary Lag Structure on Prices

True Equation and Estimation Type	a	b_0	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	c3/	k	$\bar{R}^2$	Σb_i
True Equation	1.20	-1.50	-0.50	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0		-5.0		-2.00
OLS															
No Lag	1.2109	-1.9733											-5.0724	.9916	-1.9733
1 Q Price Lag	1.1903	-1.5246	-.5530										-4.9376	.9921	-2.0776
9 Q Price Lag	1.1917	-1.5418	-.2013	-.3515	-.0070	-.2472	.4115	-.0329	-.0712	-.2723	.2440		-4.9472	.9918	-2.0607
KOYCK (implied)	1.0414	-1.7826	(-.2307)	(-.0298)	(-.0039)	(-.0005)	(-.0001)	(-0.0)	(-0.0)	(-0.0)	(-0.0)	.1294	-4.3312	.9920	-2.0475
(long-run)	(1.1962)	(-2.0475)													
ALMON - 9 Q															
2nd deg. curve	1.1914	-1.0078	-.6634	-.3809	-.1602	-.0014	.0955	.1306	.1037	.0150	-.1356		-4.9490	.9916	-2.0045
3rd deg. curve	1.1913	-1.3311	-.6272	-.1971	.0184	.0785	.0422	-.0313	-.0829	-.0534	.1163		-4.9447	.9920	-2.0676
SHILLER1/ - 9 Q															
1st deg. curve	1.1939	-1.2257	-.6538	-.2513	-.0261	.0665	.0736	-.0001	-.0695	-.0332	.0801		-4.9584	.9923	-2.0395
2nd deg. curve	1.1910	-1.3264	-.6134	-.1978	-.0027	.0624	.0523	-.0144	-.0765	-.0503	.1012		-4.9429	.9923	-2.0606
3rd deg. curve	1.1906	-1.3682	-.5763	-.2329	-.0104	.0508	.0502	-.0060	-.0769	-.0718	.1175		-4.9403	.9922	-2.1240
RIDGE2/ - 9 Q	1.2109	-1.2032	-.4541	-.2247	-.1201	-.0357	.1835	.0657	-.1445	-.0839	.1612		-5.0717	.9899	-1.8558

1/ Expressed in polynominal curve "equivalents". k = .05
2/ μ = .001
3/ Lagged dependent variable included.

GRAPH 2

Coefficient patterns for equation with rudimentary lag structure



What is somewhat disturbing is that in about half the runs the coefficient on the lagged dependent variable was found to be significant, suggesting that such a specification can produce results which confirm an erroneous hypothesis. Moreover -- also as a result of the technique -- the estimate of the current income parameter is also affected. If there were other regressors in the true relation, estimates of their coefficients would be similarly biased.

To develop this argument somewhat further, it might be borne in mind that the structure of the true equation given in Table 2 shows the Koyck technique in a relatively favorable light, considering the mis-specification. Most of the true lag coefficients are zero, and lag decay is monotonic. The constructed lag is, in fact, about the simplest imaginable. Even so, estimation results show a tendency to support an incorrect view of the structure. What would happen if this lag were lengthened and complicated a bit? To illustrate, the price lag in the Case 2 equation was extended by one quarter (to t-2) and the shape of the distribution was changed to a mild inverted V, but with the same lag sum. The equation which results from this alternate Case 2 is:

$$DV_t = -5.0 + 1.2Y_t - 0.5P_t - 1.0P_{t-1} - 0.5P_{t-2} + e_t \quad (s_e = .020)$$

When a lagged dependent variable was included in the regression, the mean estimation results were

$$DV_t = -3.6718 + .8632Y_t - 1.1858P_t + .3105DV_{t-1} \quad \bar{R}^2 = .9904$$

It is clear that a small modification of the lag process causes the Koyck-type misspecification to go to pieces in almost every imaginable way. The short-run price and income estimates are much further away from their true values than before. So also are the long-run estimates. By the usual procedure, in fact, the long run income elasticity works out to 1.2519 and the long-run price elasticity to -1.7197 (a 14 per cent error). Finally, in all but one of the runs, the coefficient on the lagged dependent variable was found to be significant; in most cases it was highly significant. There is, of course, no hope that this estimating technique can accommodate the inverted V-form of the lag.[21/]

For this reason it might be argued that the particular lag-shape chosen for this second experiment "loads the dice" against a geometric estimator. This is not necessarily the case. First, lag shapes in the real world are unknown, but since inverted V distributions have found empirical support in various areas, this kind of shape is hardly far-fetched. On these grounds alone it deserves inclusion in this experiment. Among the involved structures which may characterize real-world responses, in fact, even this form must be classed among the very simple. Secondly, for both Cases 1 and 2, the results we have chosen to tabulate are also not entirely fair to the capacities of the Almon and Shiller estimators.[22/]

21/ While not reported in detail, the Almon and Shiller estimates of this alternate equation were much more successful in picking up the change in the lag structure and adapting to the inverted V form. See Shiller /21,p.25/ for another example of an ostensibly "exponential lag" which changes shape when less restrictive priors are used.

22/ As noted above, the tabulated Almon, Shiller and ridge estimation results were deliberately and rather sharply misspecified. This was partly to give comparability to the three tables in this paper, and partly to point up the "clues" given by these more flexible estimators which can help an investigator improve the specification. Coefficient signs and significance are useful in this regard, although differences in overall equation fit may not be very helpful without elaborate testing procedures.

Returning to the basic Case 2 results, the comments which might be made about the Almon, Shiller and ridge estimates are much the same as were made with respect to the Case 1 no-lag equation. Estimation results improve as curve priors increase, with the Shiller results outperforming the Almon for equivalent curve degrees. Overall, the 2nd degree Almon traces the lag curve less well than the Koyck, but again this is an artifact of the particular curve chosen (and the conclusion reverses for the two period alternate lag case). Once again the ridge results show erratic sign changes and shifts in coefficient magnitudes; they seem generally inferior to the other estimates.

Summarizing the Case 2 findings, the evidence shows that when a single lagged term is added to only one of the basic regressors in the relation at least some of the coefficients (in various runs) on a lagged dependent variable will take on statistical significance. If the lag structure is enriched slightly, most such coefficients will turn up significant, despite the fact that the lag is still primitive (probably much less complex than in reality), and even when there is <u>no</u> lag attached to other regressors. The disturbing conclusion from these findings is that a researcher who erroneously begins with a "gap-closing" structural assumption may find his hypothesis "confirmed" by the estimation results. In some cases this may lead to serious misjudgments as to whether short-term response with respect to all variables is "elastic" or "inelastic," even if the long-term estimate is approximately correct.

We have here a situation where a false hypothesis can be upheld by the data. If this be true, then in the absence of further supportive evidence one can argue that estimation results obtained on the basis of Koyck-type hypotheses cannot really prove much at all about what the true lag-structure might be.[23/] Given this circumstance, one may ask why so many researchers continue to start with such a hypothesis and what other evidence aside from (possibly false) regression results they can produce to show that geometric adjustment exists at all? One possible form of evidence can be illustrated by turning to the Case 3 equation described above.

Case 3: Geometric lag structure on prices.

In this case it is postulated that there in fact exists (unknown to the researcher) a geometric lag-decay process on the relative price term of the hypothetical import equation. Such a process technically is of infinite duration, but beyond some point in time the influence of lagged regressors becomes negligible, and for estimation purposes it is convenient to truncate the distribution. The decay process was cut off at period t-9, so that the true function is

$$DV_t = -5.0 + 1.2Y_t + \sum_{i=o}^{9} b_o \lambda^i P_{t-i} + e_t \quad (s_e = .020) \quad b_o = -0.8, \ \lambda = 0.6$$

The sum of the b_i comes to almost -2.0 over the ten included terms.

23/ Though the arguments made here are directed specifically against the use of geometric lags, they apply also with some force against drawing firm conclusions from results obtained by using any lag structure in the absence of corroboration by yet more general forms. The low-order Almon results in both Cases 1 and 2 above are another case in point.

Obviously there are numerous possible ways an investigator who is uncertain of the true distribution might misspecify such an equation while searching for the structure. The most probable of these is that, as for some of the results shown in Tables 1 and 2, the true <u>length</u> of the lag-distribution will be guessed incorrectly. It is also probable that it will be unknown whether or not there are other lag-distributions in the equation. For purposes of simplifying the third portion of this experiment and highlighting some of the results, we have made three assumptions: 1) the investigator knows or correctly guesses that there is a lag distribution on prices, but is unsure about the income term; 2) he further surmises that the length of the price distribution is about nine quarters, but 3) he does <u>not</u> know the shape of the distribution. He therefore again applies all of the types of estimators discussed in the above sections. The findings thus obtained are summarized in Table 3 and Graph 3 on the following pages.

Results produced by the OLS estimators show the expected features. The first two equations, which understate the lag-length, also underestimate the long-run price elasticity, and somewhat overstate the income elasticity built into the structure. In the third, correctly specified equation, estimates of both these parameters come quite close to their true values, and the main problem is the erratic coefficient pattern and significance attaching to the price coefficients.

Table 3. Estimation Results for Sample Equation 3: Geometric Lag Structure on Prices ($b_i = \lambda^i b_o$; $\lambda = 0.6$, $b_o = 0.8$)

True Equation and Estimation Type	a	b_o	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9	c 3/	k	$\bar{R}^2$	Σb_i
True Equation	1.20	-.80	-.480	-.288	-.1728	-.1037	-.0623	-.0373	-.0224	-.0134	-.0081		-5.00		-1.9880
OLS															
No Lag	1.2872	-1.6268											-5.5723	.9890	-1.6268
1 Q Price Lag	1.2447	-.6652	-1.1350										-5.2944	.9909	-1.8002
9 Q Price Lag	1.1917	-.8409	-.1813	-.6395	-.1797	-.3509	-.3491	-.0702	-.0936	-.2857	.2360		-4.9472	.9920	-2.0567
KOYCK (implied	.8713	-1.2353	(-.3742)	(-.1134)	(-.0344)	(-.0104)	(-.0032)	(-.0010)	(-.0003)	(-.0001)	(-0.0)	.3030	-3.7117	.9908	-1.7723
(long-run)	(1.2500)	(-1.7723)													
ALMON - 9 Q															
2nd deg. curve	1.1837	-.7118	-.5258	-.3676	-.2372	-.1346	-.0597	-.0127	.0066	-.0019	-.0383		-4.8947	.9923	-2.0830
3rd deg. curve	1.1837	-.7691	-.5194	-.3350	-.2055	-.1204	-.0692	-.0413	-.0265	-.0140	.0064		-4.8949	.9922	-2.0940
SHILLER 1/ - 9 Q															
1st deg. curve	1.1870	-.7156	-.5241	-.3805	-.2371	-.1106	-.0205	-.0233	-.0615	-.0534	.0385		-4.9162	.9927	-2.0881
2nd deg. curve	1.1858	-.7404	-.5246	-.3602	-.2219	-.1105	-.0402	-.0306	-.0410	-.0384	.0358		-4.9084	.9925	-2.0723
3rd deg. curve	1.1891	-.7034	-.5414	-.3902	-.2336	-.0900	-.0074	-.0198	-.0850	-.0912	.0958		-4.9300	.9924	-2.0662
RIDGE 2/ - 9 Q	1.2008	-.6880	-.3806	-.4795	-.2780	-.1865	.1011	.0125	-.1368	-.0932	.1172		-4.9058	.9908	-2.0118

1/ Expressed in polynomial curve "equivalents". k = .05
2/ μ = .001
3/ Lagged dependent variable included.

Some remarks on the comparative performance of the Almon and Shiller estimates seem also to be in order. One point is that the Shiller "1st degree curve (equivalent)" estimate is produced by a method which, in a Bayesian sense, assumes that the underlying curve is really of the 1st degree, i.e., a straight line. This, too, is a misspecification of sorts. Nonetheless, over most of the lag range, this Shiller estimate outperforms an Almon of higher degree.[26/] The results using higher level curve priors underscore this conclusion, especially as regards accuracy in estimating larger, near-period coefficients.[27/] Again, the ridge estimates are clearly less stable than the others, and even seem less reliable than the OLS results.[28/]

26/ This is due to choosing a low value for k, the tightness prior. As Shiller cautions, k is not independent of measurement units.

27/ In contrast to results shown in Tables 1 and 2, in Table 3 the Shiller "3rd degree equivalent" estimates of b_o-b_2 show greater errors than those made by the lower level priors. The reasons for this are unclear, but it is probable that the anomaly would be resolved if more runs had been made.

An earlier version of this paper experimented with several values for the variance of the error term. As might be expected, the Shiller results obtained in individual runs are much more sensitive to this variance than are the Almon estimates. When the value of s_e is lowered to .002, the variance of the Shiller coefficients around their mean values is cut down greatly.

28/ Levels of μ higher than .001 once again rapidly produced a pronounced "flattening" of the b_i estimates along the abscissa. These results suggest that even searching over small values of μ is unlikely to be of much help in finding a sensible lag pattern.

VI. Conclusions

The results produced by the Case 3 experiment in this paper clearly show that several varieties of estimators are capable of approximately identifying a geometric lag decay if that is in fact the mechanism which produces a given body of data. From experiments with the more rudimentary Case 2 lag structures, we also know that using a too restrictive hypothesis may lead to the spurious conclusion that a false hypothesis is correct. In other words, Koyck-type estimation may wrongly "find" a geometric lag pattern where it does not exist, and higher level estimators are capable of finding it if it does exist. These findings tend to confirm the two propositions put forth above at the end of Section III.

We may thus return to the question posed at the outset of this paper: Under these circumstances, why is it that so many investigators continue to hypothesize that the world is governed by partial-, stock- and other types of "gap-adjustment" mechanisms which generate a Koyck-type estimator with its demonstrable deficiencies, when better methods are available? As suggested at the outset, the reason is probably that these starting points are known to lead to functions which are just plain easy to estimate.29/ This can be an important matter -- for instance,

---

29/ Some studies even appear to make errors in the way the data-transformation is carried out. Almost without exception, formal derivations of gap-type models begin in a bivariate framework, out of which pops a lagged-dependent variable. If the original relation is multivariate, however, other forms can result. A Cagan-type adaptive expectations model (gap closed at some rate on one regressor) will give an estimation equation showing lags on the other regressors. So will the original Koyck-type equation in which the lag distribution on some regressor is directly assumed to decline geometrically. Of the forms mentioned here, only Nerlove's partial-adjustment mechanism (dependent variable adjusts) can be transformed in the multivariate framework by the simple addition of a lagged dependent variable.

in the case of the kinds of trade studies cited in earlier sections. In such studies the question of short-term price and exchange-rate responsiveness is of some interest independently of the long-run considerations.[30/]

While the criticism in this paper has been largely directed at one special form of estimating lag structures, the implications are somewhat more general. The basic point is that the estimation technique derived from any hypothesis shares the limitations of that hypothesis. Nothing in this paper should be construed to suggest the author believes that geometric (or polynomial) lag processes cannot or do not really exist somewhere out there; it is only that those who purport to find them should be obliged either to a) state a strong *a priori* case in favor of such findings, or b) give some additional evidence when the shape of the lag-process is of particular interest. But in the vast mass of published empirical conclusions on the shape of distributed lags, it is difficult to find a study in which an author actually backs up, in either of these two fashions, an assertion that a lag has particular properties. One result is that the literature is filled with masses of seemingly contradictory results. The "checking" process suggested in this paper has obvious limitations of its own, but it is at least apparent that such verification can help clarify what conclusions the data do or do not sustain.

30/ One thinks, for instance, of the chagrin evinced by most forecasters when the exchange rate changes set in December, 1971, under the Smithsonian Agreement failed to produce their desired results after a few quarters. If, as Table 3 suggests, geometric estimates overstate short-term responsiveness, this is one possible reason.

APPENDIX: Correlation Matrix for Sample Data

	DV	Y	P_o	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9
DV	1.0											
Y	.990	1.0										
P_o	-.879	-.821	1.0									
P_1	-.887	-.841	.950	1.0								
P_2	-.891	-.849	.909	.958	1.0							
P_3	-.890	-.855	.867	.921	.961	1.0						
P_4	-.878	-.860	.800	.874	.923	.962	1.0					
P_5	-.863	-.857	.756	.817	.879	.925	.965	1.0				
P_6	-.855	-.858	.709	.772	.825	.883	.929	.968	1.0			
P_7	-.850	-.859	.690	.741	.790	.837	.890	.930	.969	1.0		
P_8	-.848	-.865	.669	.727	.764	.806	.845	.893	.931	.970	1.0	
P_9	-.848	-.874	.653	.712	.754	.784	.815	.848	.894	.934	.973	1.0

NOTE: Dependent variable (DV) includes random error component.

References

1. Almon, Shirley. "The Distributed Lag Between Capital Appropriations and Expenditures," Econometrica, January, 1965, pp. 178-196.

2. Branson, William H. "A Disaggregated Model of the U.S. Balance of Trade," Staff Economic Studies #44, Board of Governors of the Federal Reserve System, February, 1968.

3. Cagan, P. "The Monetary Dynamics of Hyper Inflations," in M. Friedman, ed., Studies in the Quantity Theory of Money (Chicago: University of Chicago Press), 1956.

4. Corradi, C. and Gambetta, G. "The Estimation of Distributed Lags by Spline Functions" (processed), 1975.

5. Dhrymes, Phoebus J. Distributed Lags: Problems of Estimation and Formulation, (San Francisco: Holden-Day), 1971.

6. Goldstein, Morris, and Khan, Mohsin S. "Large Versus Small Price Changes and the Demand for Imports," (mimeographed) International Monetary Fund, DM/74/20, February 27, 1975.

7. Gregory, R.G. "United States Imports and Internal Pressure of Demand: 1948-68," American Economic Review, March, 1971, pp. 27-47.

8. Griliches, Zvi. "Distributed Lags: A Survey," Econometrica, January, 1967, pp. 16-49.

9. Hoerl, Arthur E., and Kennard, Robert W. "Ridge Regression: Applications to Nonorthogonal Problems," Technometrics, February, 1970, pp. 69-82.

10. ______________________________. "Ridge Regression: Biased Estimation for Nonorthogonal Problems," Technometrics, February, 1970, pp. 55-67.

11. Hooper, Peter III. "An Analysis of Aggregation Error in U.S. Merchandise Trade Equations," Ph.D. Dissertation in Economics, University of Michigan, 1974.

12. Koyck, L.M. Distributed Lags and Investment Analysis (Amsterdam: North Holland), 1954.

13. Kwack, Sung Y. "The Determination of U.S. Imports and Exports: A Disaggregated Quarterly Model, 1960.III - 1967.IV," Southern Economic Journal, January 1972, pp. 302-15.

14. Maddala, G.S. "Ridge Estimators for Distributed Lag Models," NBER Working Paper No. 69 (processed), October, 1974.

15. Miller, Joseph C. and Fratianni, Michele. "The Lagged Adjustment of U.S. Trade to Prices and Income," Journal of Economics and Business, Spring, 1974, pp. 191-198.

16. Nerlove, M. Distributed Lags and Demand Analysis, USDA, Agriculture Handbook No. 141, Washington, 1958.

17. Prachowny, Martin F. A Structural Model of the U.S. Balance of Payments (Amsterdam: North Holland), 1969.

18. Rao, Sista V. "An Empirical Study of U.S. Foreign Trade Sector According to Standard International Trade Classification During 1950.I - 1970.IV," Ph.D. Dissertation in Economics, University of Pennsylvania, 1971.

19. Rappoport, Paul N. "Least Squares and its Alternatives in the Estimation of Dynamic Economic Models" (processed), 1974.

20. Schmidt, Peter. "A Modification of the Almon Distributed Lag," Journal of the American Statistical Association, September, 1974, pp. 679-681.

21. Shiller, Robert J. "Alternative Prior Representations of 'Smoothness' for Distributed Lag Estimation," NBER Working Paper No. 89 (processed), June, 1975.

22. ______________________________. "A Distributed Lag Estimator Derived from Smoothness Priors," Econometrica, July, 1973, pp. 775-787.

23. ______________________________. "Estimation of the Investment and Price Equations of a Macroeconometric Model," Staff Economic Studies No. 61, Board of Governors of the Federal Reserve System, 1971.

24. Taylor, William E. "Smoothness Priors and Stochastic Prior Restrictions in Distributed Lag Models," International Economic Review, October, 1974, pp. 803-804.

25. Wilson, John F. "Yet Another Econometric Model of U.S. Imports, 1958-1971, Disaggregated by End-Use Commodity Groups and Region of Origin," Ph.D. Dissertation in Economics, University of Pennsylvania, 1974.