STAFF PAPERS IN ECONOMICS

Agricultural Economics & Economics Dept.

College of Agriculture
College of Letters and Science

AAEA paper 1978

Montana State University, Bozeman

NONSTOCHASTIC DIFFERENCE EQUATIONS, DISTRIBUTED LAGS, AND AGRICULTURAL SUPPLY

Oscar R. Burt

Staff Paper 78-9

Paper presented at AAEA meeting, Blacksburg, Va., Aug. 6-9, 1978.

NONSTOCHASTIC DIFFERENCE EQUATIONS, DISTRIBUTED LAGS, AND AGRICULTURAL SUPPLY

by

OSCAR R. BURT

Montana State University

I. Introduction

One unequivocal aspect of agricultural supply response is its complex dynamic structure. The two decades since Nerlove's pioneering work have given us a variety of dynamic models and statistical estimation procedures. Nearly all of the research has focused on various forms of distributed lags in response, and statistical methods for estimating the parameters involved (see Dhrymes, 1971a, for many of these models). A recent survey of agricultural supply response studies which used the "Nerlove Model" and various extensions of it contained 190 references to the literature (Askari and Cummings), a good indication of the practical importance of such models.

Usually a dynamic economic problem would suggest differential or difference equations among the endogenous variables, but many of the distributed lag behavioral assumptions in supply models are only dynamic in the structure expressed through lagged values of the exogenous variables. An example is the adaptive expectations model of Nerlove, which reduces to a linear equation in historic prices with the coefficients on lagged prices declining geometrically with respect to time measured backwards from the present. Models which include only lagged exogenous variables would be satisfactory if the set of current and lagged independent variables were complete, the history of the data sufficiently long, and the structure constant over the sample period. Under this idealized situation, and assuming a correct specification of the supply equation, any lagged values of the dependent variable would be redundant because the history of the independent variables provides complete information which cannot be improved upon.

Then we have other models, such as Nerlove's partial adjustment model, which directly imply a difference equation in the dependent variable. The rationale for the partial adjustment model was implicit in Cassels and centers on the time required for output changes of the firm. Existing assets of the firm (industry), currently employed labor, and other short-run conditions limit the speed of economic adjustment to changed prices. Nevertheless, an idealized specification of the supply equation with all relevant exogenous variables lagged over a sufficiently long history of data would capture all of these short-run conditions and make the difference equation unnecessary.

The problems with a model which relies on such a complete specification and historical data series that all relevant information is taken into account are quite obvious. First, measurements on some variables needed in the complete specification will nearly always be missing. Second, the number of parameters to be estimated would be very large, creating a problem in degrees of freedom. The practice of smoothing the coefficients on lagged values of the exogenous variables by specifying a polynomial or frequency distribution curve would partially solve this difficulty (see for example Baritelle and Price; Chen, Courtney, and Schmitz; and Kulshreshtha).

Third, the structure of supply response cannot be expected to remain unchanged for very long periods, even with a tolerable level of specification error. Therefore, the limited historical period over which the model is appropriate will severely restrict the information set implicit in the model.

Fourth, the correct specification is almost certainly nonlinear in the lagged values of the exogenous variables. The usual arguement that linear relationships are adequate as an approximation is less convincing when we are dependent on the correct specification over an extended historical period to capture all the rigidities created by past investments in the productive capacity of the industry.

Specification of a dynamic behavioral equation, such as for agricultural supply, is quite analogous to choosing the set of state variables which are used in a dynamic optimization problem. The state variables must be few in number but capture most of the information about the history of the decision process. In supply estimation, the predetermined variables must be efficient at summarizing information on the history of the many exogenous variables and the functional form in which they influence supply response.

It would appear that a recent history of the levels of the supply response variable (planted acres, for example) would summarize a large part of the information implied by the historical time series of the exogenous variables which jointly describe the dynamic structure of supply. Nerlove's partial adjustment model (Nerlove, 1958, p. 62) is a simplified case where the level of response last year summarizes the entire history. This model is easily generalized to higher order lags in the adjustment process, yielding correspondingly higher order difference equations and lags on the price variable. This kind of generalization is capable of handling the rigidities in production which typically distinguish various lengths-of-run.

A difference equation in the supply response variable can be justified directly without recourse to the partial adjustment argument. With a first order difference equation, we could interpret the response level last year as indicative of current capacity for aggregate production of the commodity. In a second order equation, we would have an indirect measure of capacity for the last two years, or alternatively, we could interpret the second order equation as reflecting the level of capacity last year and the change in capacity last year. A third order equation would also reflect the rate of change in capacity to produce. We are using the term "capacity" rather loosely here; the same interpretation of the difference equation can be made with respect to inertia, habit, tradition, and crop rotations. In applications to livestock inventories the vague notion of capacity would have special connotations, likewise for perennial crops such as tree crops.

The other main source of dynamic behavior in supply is associated with the formation of price expectations. The usual assumption is that of geometrically declining weights on experienced prices backwards in time, and after a suitable transformation, this specification can be reduced to a first order difference equation in the supply response variable and only last year's price remains as an independent variable (Nerlove, 1956, p.502). If a difference equation has already been specified in the response variable before the transformation, the difference equation is of one higher degree after the transformation (Nerlove, 1958, p.64).

Even when Nerlove's simple adaptive price expectations model is appropriate and the above transformation has been made, it might be necessary to include second or higher order lagged prices explicitly in the response equation. These extra lagged prices taken jointly with the lagged values of the response variable can capture information on rigidities in the industry that would fall into the error term otherwise. The higher the level of aggregation, the more likely it is that these lagged price variables will reflect extra information with respect to restrictions on production response, i.e., information beyond that provided by lagged values of the response variable conjunctively with an exogenous variable at a single point in time.

It is concluded that a combination of many economic, technological, and behavioral factors and constraints implies a difference equation in the supply response variable, and in addition, moderately small order lags on the exogenous variables are likely to be required to get a good approximation of the supply response equation.

Other justifications for this type of specification are to be found in the economics literature on distributed lags, such as the general result of Jorgenson that any lag distribution can be approximated by a rational distributed lag function, or that the distributed lag pattern can be approximated by the Pascal distribution (Solow).[1/] Then going back another step in the logical process, Grether has recently shown that a rational distributed lag follows from fairly general assumptions about the underlying economic process which gives rise to an unobservable variable in a regression equation.

The next section discusses problems of irreversibilities in agricultural supply studies, and the third section considers identification of various components in the dynamic structure of supply response equation. Section IV distinquishes between stochastic and nonstochastic difference equation; section V takes up statistical estimation of dynamic regression equations and section VI discusses the interpretation and specification of the error term. Some examples of applications of the methodology are given in section VII along with suggestions on the analysis of results from dynamic models of supply. The final section summarizes the main results and conclusions on supply estimation from time series data.

II. Irreversibilities in Supply

The notion of irreversibilities in supply response have been with us for a long time (Cassels), but the practice of statistically testing for irreversibility has received considerable attention recently (Houck; Traill, Coleman, and Young; and Wolffram). It would appear that finding irreversibility in a linear supply function with respect to price changes is a symptom of specification error associated with the dynamic structure of the model. In fact, an irreversibility with respect to one variable in the supply response function would seem to suggest irreversibility in most other variables of the equation.

A sufficiently high order difference equation in the supply response variable jointly with a relatively short distributed lag on price, and possible other exogenous variables, should remove the irreversibility problem. A first order difference equation in annual data will reflect the level of output last year; a second order equation will indicate whether output was increasing or decreasing during the last two years; while a third order equation can also measure rate of change in output. This capability of difference equations taken jointly with second or third order lagged prices should take account of essentially all the factors associated with irreversibilities. Price level last year, its change during the last two years, and possibly the rate of change during the last three years taken jointly with the same type of information on output would be expected to capture the dynamics of supply response commonly manifested as irreversibilities in static linear regression models. Of course, we would expect something as simple as a second order difference equation to be adequate in many applications.

The use of relatively high order lags on price, and sometimes other exogenous variables, without using a difference equation in the supply response variable has been used quite effectively in some studies; examples are Kulshreshtha for slaughtered beef and Baritelle and Price for tree crops. This success of econometric models with implicit irreversibilities without using difference equations merely raises the question of relative performance of two types of dynamic models. The author leans heavily in favor of the difference equation models, particularly nonstochastic difference equations to be introduced in section IV, reasons for this preference were covered in the introductory discussion. An example of the superiority of difference equations will be given later using acreage response of U.S. wheat producers.

III. Identification Problems[2/]

Detailed specifications of how producers' price expectations are formulated and the adjustment mechanism that governs their short-run changes in output would be viewed by many econometricians as superior to simply specifying a second or third order difference equation in the response variable with second or third order lagged prices as independent variables. The latter relatively loose specification would seem to border on mere empiricism -- measurement without adequate theory. Let us examine the virtues and limitations of these two approaches to specification or model building.

Detailed Approach

One of the goals in econometric work is to test hypotheses, and a detailed specification provides much sharper hypotheses to be tested. There is no a priori reason to expect that individual parameters which describe the various components of the dynamic structure of supply response can be identified from economic time series data. Nerlove encountered a situation where two key parameters were not identified when he combined the partial adjustment and adaptive price expectations hypotheses into one model (Nerlove 1958, p.64). Often when the parameters can be identified in principle, a serious problem of statistical precision will emerge when the model is fitted to time series data (Nerlove 1958, chapter 9 and Behrman).

Most hypotheses about the structure of supply response come from the theory of the firm and postulated behavior of an individual decision agent, but the empirical equation is fitted to aggregate time series data (regional or national), which raises many questions about compatability between an hypothesis and the data used to test it. The nature of these problems is illustrated with acreage response of U.S. wheat producers.

Land resources used in wheat production are extremely heterogenous, varying from marginal land for crop production in The Great Plains to highly productive land in the corn belt, as well as irrigated areas throughout the country. Although we might expect producers' formulations of price expectations to be quite similar across the nation, the rigidities and constraints on changes in acreage which characterize short-run adjustments are much different among the various regions of production. If new cropland is brought into production in the semi-arid Great Plains, the land is summer-fallowed the first year to build a soil moisture reserve and allow the sod to decay, which implies a technical constraint on short-run acreage expansion. On the other hand, shifting from one annual crop to another on existing cropland can be effected easily from year to year, except in parts of The Northern Great Plains where it is difficult to get winter wheat planted behind another crop because of the short growing season. But another constraint exists in The Great Plains in that the traditional dryland farming practice is to summer-fallow the cropland every other year. This tradition is not completely rigid since there is a transition area in going from east to west in The Great Plains where summer-fallow is a marginal practice which is used when prices are relatively low but abandoned to some degree when prices are high. These rigidities associated with crop rotations are not limited to The Great Plains either; winter wheat cannot usually be planted after corn harvest in the fall which causes a constraint on rotation changes in The Corn Belt. Then there are also the constraints associated with specialized farm equipment for individual crops.

These technical agronomic and practical economic constraints on changes in wheat acreage will tend to confound the test of a specific hypothesis such as Nerlove's partial adjustment model. There is a possibility of incorporating a sufficient number of concomitant variables into the regression equation to remove influences of the technical rigidities, but the author would not be optimistic about this being successful.

Many different detailed specifications on the forms of the lag distributions on independent variables would be about equally plausible on a priori grounds, as would many varied specifications of the short-run adjustment process. But experience with time series data and the usual size samples leads one to expect that several detailed specifications are likely to fit the data about equally well in supply estimation. This is nothing extraordinary in econometric research and merely a manifestation of the general indentification problem of science (Russell, p.330):

> "Every finite set of observations is compatible with a number of mutually inconsistent laws, all of which have exactly the same inductive evidence in their favour."

It is this author's opinion that overly specific models in supply response studies have a tendency to delude us into thinking that economic measurement is more precise than it is.

A Robust Specification

What do we lose by merely assuming that a relatively low order difference equation in conjunction with modest lags on the exogenous variables is capable of describing the dynamics of agricultural supply response? We give up an opportunity to identify the individual components which comprise the underlying forces behind a dynamic model of supply response. The mechanisms describing the formulation of price expectations and the rigidities in short-run adustments are confounded, along with any other dynamic factors such as crop rotation constraints.

An advantage of the more general specification would appear to be less frequent specification error, or at least, the nature of the specification error would tend to be that which the data is incabable of detecting. The detailed approach would in principle require the testing of many different algebraic forms for the separate dynamic components of the model, while the general approach would involve sequential testing of various orders of lags on the variables. In practice, the detailed approach is likely to be applied without testing various algebraic forms on the dynamic components of the model.

A primary justification for the general difference equation specification is the result of Jorgenson that any distributed lag function can be approximated by a rational lag function. The difference equation emanates from multiplication of both sides of the response equation by the denominator of the rational lag function, which is a polynomial in the lag operator.

In agricultural supply estimation we also have a rather direct basis for the difference equation which emanates from rigidities in short-run adjustments. One way to view a relatively low order difference equation in supply response with several exogenous variables is that we force the same denominator on the rational lag function associated with each exogenous variable, and also, use this denominator function to describe the phenomena in supply response which warrant a difference equation directly. Ultimately the worth of this methodology must be determined by experience in application, some limited results of which are presented in section VI.

IV. Stochastic Versus Nonstochastic Difference Equations

Problems of statistical estimation are postponed so that we can contrast the conceptual aspects of stochastic and nonstochastic difference equations. We use Nerlove's partial adjustment model to illustrate the conceptual framework (Nerlove 1958).

The Partial Adjustment Model

Let x_t^* be desired long-run equilibrium production of a crop and let x_t be observed production in year t, while price is denoted by p_t. Long run supply is given by

$$(1) \quad x_t^* = a + bp_{t-1} \ .$$

The supply adjustment equation is

$$(2) \quad x_t - x_{t-1} = \gamma(x_t^* - x_{t-1}) \ , \ 0 < \gamma \leq 1 \ .$$

The supply response variable has been taken as production instead of acreage to make the role of the error term more obvious. Clearly, x_t^* is a conceptual variable which is unobservable. Should the supply adjustment equation be (2) or should it be redefined in expectational form as

(3) $E(x_t) - E(x_{t-1}) = \gamma[x_t^* - E(x_{t-1})]$,

where $E(\cdot)$ denotes the expectation operator?

If (2) is assumed, the size of the adjustment will depend heavily on weather conditions in year t-1 since a high yield and the associated large production will tend to reduce x_t and vice versa. Does this make sense? The concept of a partial adjustment model would appear to fit the nonstochastic model of (3) better than (2). The expected values of production are interpreted as quantities produced without any unusual events or conditions, given complete information implicit in the model defined by (1) and (3) jointly. Substitution of (1) into (3) yields

(4) $E(x_t) = (\gamma a) + (\gamma b)p_{t-1} + (1 - \gamma) E(x_{t-1})$,

and the right hand side makes it clear what the conditional information is upon which $E(x_t)$ depends. Solution of the difference equation of (4) shows that $E(x_t)$ depends on the entire history of prices.

A statistical equation is obtained by adding a disturbance term to each side of (4) to get

(5) $x_t = (\gamma a) + (\gamma b)p_{t-1} + (1 - \gamma) E(x_{t-1}) + u_t$.

This direct approach for getting (5) from (4) assumes that x_t^* is nonstochastic and uses the identity $x_t = E(x_t) + u_t$, where $E(u_t) = 0$.[3/]

Let us consider (5) from a rather direct, even empirical, point of view. The term $E(x_{t-1})$ is an indirect measure of the inertia for growing the crop and reflects such things as investments in specialized equipment, established rotations, existing labor supply and farm organizations, and all the factors which tend to create short-run constraints on production. If random factors subsumed in u_{t-1} have an effect on production the following year, their impact is probably different, maybe even opposite in sign, than the impact of $E(x_{t-1})$. A generalization of (5) to accomodate the possible influence of a lagged disturbance term would be

(6) $x_t = (\gamma a) + (\gamma b)p_{t-1} + (1 - \gamma)\ E(x_{t-1}) + \theta u_{t-1} + u_t$,

which will be recognized as an equation with a moving average error term.

Suppose that the actual adjustment process is as given by (3), but the operational regression equation is derived by using (2). Direct substitution of (1) into (2) gives

(5)´ $x_t = (\gamma a) + (\gamma b)p_{t-1} + (1 - \gamma)x_{t-1} + v_t$,

where v_t has been added as a disturbance term which emanates from (1) and/or (2) having a disturbance term added. But $x_{t-1} = E(x_{t-1}) + v_{t-1}$

which substituted into (5)´ yields

(6)´ $x_t = (\gamma a)+(\gamma b)p_{t-1} + (1 - \gamma)E(x_{t-1}) + (1 - \gamma)v_{t-1} + v_t$

Comparison of (6)´ with (6) reveals that the use of (5)´ will give the same results as (6) if and only if $\theta = 1 - \gamma$, a highly unlikely event.[4/]

If the correct adjustment constraint is given by (3), then the use of (2) and the implied estimation equation in (5)´ results in an "errors in variables" model. In Griliches' study of the aggregate U.S. farm supply function, he was puzzled by the relative instability of the distributed lag model. One of the possible causes suggested by Griliches was (Griliches 1960, p.291):

> "the fact that measured output is not necessarily equal to planned output, due to "weather" and other random effects. This last factor would lead to a downward bias in the estimate of the coefficient of lagged output since the adjustment assumed by the model proceeds from the previously "planned" output, of which actual output is not an error-free measure. The presence of random measurement errors in an "independent" variable usually leads to a downward bias in its estimated coefficient."

The above point is particularly plausible in view of the fact that relatively better results were obtained by Griliches for aggregate supply response of livestock and livestock products than for crops or total aggregate supply, the latter two being more susceptible to random weather factors. A weather index was included as an independent variable in the regressions for total supply and the crops aggregate, but such an index can only partially accomplish the needed adjustment for weather.

When the variable x_t is taken as acreage instead of production, the same considerations apply with respect to the random and systematic components of the variable, but the nature of the random component is not so obvious. Instead of the disturbance u_t being dominated by variations in crop yields and the consequent variation in production, it will be the sum of all those influences on acreage which cannot be explicitly incorporated into the acreage response equation. Whether a generalization of the error term such as in (6) is needed is an empirical question, but it seems very unlikely that the random and systematic components of x_{t-1} would have the same effect on x_t.

Nonstochastic Difference Equations

Let us begin with the Nerlove first order difference equation which has been traditionally fitted by least squares, and to simplify things, let there be only one exogenous variable:

(7) $y_t = \alpha + \beta p_{t-1} + \lambda y_{t-1} + u_t$

The above difference equation is "stochastic" by two criteria, (1) the random error u_t, and (2) the lagged dependent variable y_{t-1} enters as an explanatory variable. The distinction made in this article between stochastic and nonstochastic difference equations is on the basis of whether y_{t-1} or $E(y_{t-1})$ enters the equation as the predetermined variable.

If we take expectations of both sides of (7), the result is

(8) $E(y_t) = \alpha + \beta p_{t-1} + \lambda y_{t-1}$

since the usual assumption is made that $E(u_t) = 0$. Technically (8) is the conditional expectation of y_t, given the observed value of y_{t-1}. This is consistent with the way (7) would be fitted statistically, where y_{t-1} is simply a second independent variable in the regression equation.

The counterpart of (8) for a first order nonstochastic difference equation is

(9) $E(y_t) = \alpha + \beta p_{t-1} + \lambda E(y_{t-1})$

which defines the unconditional expectation of y_t. If we iterate (9) by successive substitutions for $E(y_{t-j})$, the result is

$$
\begin{aligned}
E(y_t) &= \alpha + \beta p_{t-1} + \lambda[\alpha + \beta p_{t-2} + \lambda E(y_{t-2})] \\
&= \alpha(1 + \lambda) + \beta(p_{t-1} + \lambda p_{t-2}) + \lambda^2 E(y_{t-2}) \\
&\;\;\vdots \\
&= \alpha(1 + \lambda + \lambda^2 + \ldots) + \beta(p_{t-1} + \lambda p_{t-2} + \lambda^2 p_{t-3} + \ldots)
\end{aligned}
\tag{10}
$$

which shows that the unconditional expectation of y_t is dependent only on the historical series of the exogenous variable. In contrast, the conditional expectation of y_t depends on both the historical series of the exogenous variable and the dependent variable itself.

If we truncate the iterations in (10) when $\lambda^t E(y_o)$ appears, the result is

(11) $$E(y_t) = \alpha(1 + \lambda + \ldots + \lambda^{t-1}) + \beta(p_{t-1} + \lambda p_{t-2} + \ldots + \lambda^{t-1} p_o) + \lambda^t E(y_o)$$

If the parameters α, β, and λ were known together with $E(y_o)$, $E(y_t)$ could be easily calculated for $t = 1, 2, \ldots T$, where T is the sample size.

V. Statistical Estimation

Adding a disturbance term to both sides of (9) gives

(12) $y_t = \alpha + \beta p_{t-1} + \lambda E(y_{t-1}) + u_t$

since $y_t = E(y_t) + u_t$. Under classical assumption on the disturbance u_t, maximum likelihood estimates of the unknown parameters can be obtained by searching over various values of λ while treating $p_{t-1} + \lambda p_{t-2} + \ldots + \lambda^{t-1} p_o$ and λ^{t-1} as independent variables for given λ. A least squares linear regression can then be fitted with $E(y_o)$ conceived as an unknown parameter, which is obvious from (11). This approach to estimation of the geometrically distributed lag on one independent variable was first recognized by Klein, but an interpretation of the model as a nonstochastic difference equation seems to have been largely overlooked.[5/] The method is not much different than Nerlove's iterative method (Nerlove 1958); the essential difference is that Nerlove's method estimates $E(y_o)$ by applying the geometric lag to pre-sample values of the independent variable instead of treating $E(y_o)$ as a parameter.[6/]

Extending (9) to a second order difference equation

(13) $E(y_t) = \alpha + \beta p_{t-1} + \lambda E(y_{t-1}) + \mu E(y_{t-2})$

and attempting an explicit solution to get the counterpart of (11) will demonstrate the practical difficulties of using a direct search procedure analogous to the first order case. Taking $E(y_o)$ and $E(y_{-1})$ as given values and iterating (13) from $t = 1$, the general expression for observation t, which is the counterpart of (11), will contain a weighted sum of the lagged values of the exogenous variable with the weights being functions of λ and μ , and there will be a term for each $E(y_o)$ and $E(y_{-1})$ with a respective coefficient involving λ, μ, and the observation number t. For given λ and μ, this explicit solution can be viewed as a linear regression with $E(y_o)$ and $E(y_{-1})$ treated as unknown parameters. Therefore, a direct search on λ and μ combined with many solutions of a linear regression could be used to obtain least squares estimates of the unknown parameters (maximum likelihood if normality is assumed on the disturbance term).

Not only are the computations quite heavy for anything higher than a first order equation, but we will probably want to consider general specifications on the error term such as serial correlation, moving average, or a combination of both. Therefore, a general nonlinear least squares algorithm was developed to handle essentially any order of difference equation jointly with serial correlation/moving average error terms.

Error Term with Classical Properties

We begin with the simplest case of a first order difference equation and the classical assumptions on the error term so that the reader can readily see the relationship of the method to the direct search procedure. Thus our model is given by (12) under the assumption that $E(u_t) = 0$ and $E(u_t^2) = \sigma^2$ for $t = 1, 2, \ldots T$, and $E(u_i u_j) = 0$, $i \neq j$.

In order to apply a nonlinear least squares algorithm, we must be able to calculate $E(y_t)$, $t = 0, 1, 2, \ldots T - 1$ and the partial derivative of y_t with respect to each parameter for $t = 1, 2, \ldots T$, where T is the number of data points in the sample (Draper and Smith). Some initial estimate of the parameters is assumed; in practice, these estimates will usually be calculated by using y_{t-1} in place of $E(y_{t-1})$ and fitting a linear regression equation. Let the transitory estimates at a particular iteration be denoted by $\tilde{\alpha}$, $\tilde{\beta}$, and $\tilde{\lambda}$. We take y_o as an estimate of $E(y_o)$; more will be said about this later.

At a given iteration, the estimate of $E(y_t)$ is given by

(14) $\tilde{y}_t = \tilde{\alpha} + \tilde{\beta} p_{t-1} + \tilde{\lambda}\tilde{y}_{t-1}$,

which can be calculated recursively with $\tilde{y}_o$ set equal to y_o. Partial derivatives of $\tilde{y}_t$ with respect to each parameter estimate are also calculated recursively, with the initial condition given by zero since $\tilde{y}_o$ is taken as a constant. Let us illustrate with $\tilde{\beta}$,

$$\partial\tilde{y}_o/\partial\tilde{\beta} = 0$$

$$\partial\tilde{y}_1/\partial\tilde{\beta} = p_o + \tilde{\lambda}(\partial\tilde{y}_o/\partial\tilde{\beta}) = p_o$$

$$\partial\tilde{y}_2/\partial\tilde{\beta} = p_1 + \tilde{\lambda}(\partial\tilde{y}_1/\partial\tilde{\beta})$$

$$\vdots$$

(15) $\partial\tilde{y}_t/\partial\tilde{\beta} = p_{t-1} + \tilde{\lambda}(\partial\tilde{y}_{t-1}/\partial\tilde{\beta})$.

Since (11) is an explicit solution of the difference equation, $\tilde{y}_t$ calculated recursively from (14) will be the same as if $\tilde{\alpha}$, $\tilde{\beta}$, and $\tilde{\lambda}$ were substituted for the unknown parameters in (11) and $E(y_o)$ set equal to y_o. Likewise, the recursive calculation of the partial derivative for $\tilde{\beta}$ in (15) will give the same results as if (11) were used. From this observation, it is clear that the nonlinear least squares algorithm could just as well be applied to the explicit solution of the difference equation and used in place of the search procedure except that the above method does not estimate a parameter for $E(y_o)$. But there is no reason why this additional parameter cannot be estimated simultaneously with α, β, and λ . Basically all we need is another column vector of partial derivatives. Let the unknown parameter for $E(y_o)$ be η ; then $\partial\tilde{y}_t/\partial\tilde{\eta} = \tilde{\lambda}^t$ which is obvious from (11) or a recursive calculation such as (15).

The use of y_o as an estimate of $E(y_o)$ is justified for large samples by the fact that the least squares estimates are consistent and efficient regardless of the value used for $E(y_o)$ (Dhrymes 1971b.). Nevertheless, a rather absurd value such as zero for $E(y_o)$ could produce poor estimates in small samples, but the results of Pesaran suggest that replacing $E(y_o)$ by y_o should give good results compared to estimation of $E(y_o)$ as a parameter. Only Monte Carlo studies can provide additional information on this issue. It would seem that there could be some advantage in smoothing a short series of the presample values of the dependent variable to get a more representative estimate of $E(y_o)$ than simply y_o itself, particularly if the dependent variable has inherently large variation around its mean such as would be the case with crop yield measurement.

One of the main virtues of the nonlinear least squares procedure just described is its ease of generalization to essentially any order of nonstochastic difference equation.[7/] We merely note that the computational burden is highly dependent on the number of observations since the iterative calculation of partial derivatives for the parameters is rather time consuming, but this is certainly no problem with annual time series data frequantly used in agricultural supply estimation. Also, the number of parameters has a large influence on computational time for two reasons, (1) the usual reason that the matrix dimensions become large, and (2) the number of partial derivatives which must be calculated is equal to the number of parameters.

The serious problem of getting initial starting values for the parameters when using nonlinear least squares is of minor importance in this case because we can use a linear regression with $E(y_{t-j})$ replaced by y_{t-j} to get these initial estimates. Except in cases where the variance of the residuals is quite large such as with crop yield data, we would expect this initial linear regression to put us in the neighborhood of the least squares solution. The author has encountered little difficulty even with U.S. wheat yield as the dependent variable.

Serially Correlated Error Term

The linear nonstochastic difference equation is now generalized to allow serial correlation in the disturbance term. For discussion purposes we use a second order difference equation and second order serial correlation; the extension to higher order systems is then obvious. Let there be two exogenous variables p_{t-1} and z_t, then the model is

$$(16) \quad y_t = \beta_1 + \beta_2 p_{t-1} + \beta_3 z_t + \lambda_1 E(y_{t-1}) + \lambda_2 E(y_{t-2}) + u_t,$$

$$(17) \quad u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \varepsilon_t,$$

$$E(u_t) = E(\varepsilon_t) = 0, \ t = 1, 2, \ldots T,$$

$$E(\varepsilon_i \varepsilon_j) = 0 \ , \ i \neq j$$

$$= \sigma^2, \ i = j, \text{ all } i \text{ and } j \ .$$

Since (16) can in principle be reduced to an equivalent equation with the unobservable variables $E(y_{t-1})$ and $E(y_{t-2})$ replaced by a nonlinear function of the unknown parameters and exogenous variables, as was illustrated for the simpler case of (12), we can treat (16) and (17) as a model where $E(y_t)$ is nonlinear in the parameters and the disturbance is described by a second order serial correlation process. The usual autoregressive transformation can be applied by lagging (16) one and two periods, multiplication by ρ_1 and ρ_2 respectively, and subtracting the two results from (16) to get

$$(18) \quad y_t = \beta_1(1 - \rho_1 - \rho_2) + \beta_2(p_{t-1} - \rho_1 p_{t-2} - \rho_2 p_{t-3}) + \beta_3(z_t - \rho_1 z_{t-1} - \rho_2 z_{t-2}) + \lambda_1[E(y_{t-1}) - \rho_1 E(y_{t-2}) - \rho_2 E(y_{t-3})] + \lambda_2[E(y_{t-2}) - \rho_1 E(y_{t-3}) - \rho_2 E(y_{t-4})] + \rho_1 y_{t-1} + \rho_2 y_{t-2} + \varepsilon_t,$$

$$t = 3, 4, \ldots T \ .$$

Since the new residual ε_t has the classical properties, least squares estimates of the parameters in (18) will be maximum likelihood if we assume ε_t is normal and treat the presample values y_2, y_1, $E(y_o)$, and $E(y_{-1})$ as fixed from sample to sample,[8/] or estimate $E(y_o)$ and $E(y_{-1})$ as additional parameters.

Based on the results of Pesaran, we would appear to lose very little by estimation of $E(y_o)$ and $E(y_{-1})$ by y_o and y_{-1}. Therefore, we have a relatively simple procedure to apply and program for an electronic computer. Initial estimates for all parameters except ρ_1 and ρ_2 can be obtained with a linear regression where y_{t-j} replaces $E(y_{t-j})$, and then the residuals from that linear regression can be used to get initial estimates of ρ_1 and ρ_2. Since the parameter estimates for a stochastic difference equation are not consistent when ordinary least squares is used in the presence of serial correlation in the residuals, care should be taken to try several different starting positions with respect to ρ_1 and ρ_2. Problems of local minima in the residual sum of squares should not be taken lightly in nonlinear estimation.

Joint Serial Correlation and Moving Average Error

A general specification of the n^{th} order nonstochastic difference equation with an additive random disturbance which follows a joint serial correlation/moving average process is

$$(19) \quad y_t = \beta_o + \beta_1 z_{1t} + \ldots + \beta_k z_{kt} + \lambda_1 E(y_{t-1}) + \ldots + \lambda_n E(y_{t-n}) + u_t$$

$$(20) \quad u_t = \rho_1 u_{t-1} + \ldots + \rho_m u_{t-m} - \theta_1 \varepsilon_{t-1} - \ldots - \theta_q \varepsilon_{t-q} + \varepsilon_t$$

where ε_t has the classic properties outlined under (17). The order of serial correlation is m and the moving average is of order q. The signs on the moving average parameters, $\theta_1, \ldots \theta_q$, are negative by the usual convention in the time series literature (Box and Jenkins, Nelson).

The general autoregressive transformation, illustrated by the derivation of (18) from (16) and (17) for second order serial correlation, can be applied to (19) and (20). The result is an m^{th} order generalization of (18) with respect to the serial correlation parameters, and also, the following moving average error terms are added,

(21) $- \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q}$.

Thus it is seen that the resulting equation is of the same general form as (18) except for the presence of a moving average error term.

All of the parameters can be estimated simultaneously by nonlinear least squares. Note that the terms involving the moving average error in (21) constitute a linear difference equation in the $\{\varepsilon_t\}$ so that the method explained in the last section for the difference equation in $E(y_t)$ can be applied. Initial values for ε_o, ε_{-1}, . . . ε_{-q+1} must be assumed or else these initial values can be treated as additional parameters to be estimated, just like $E(y_o)$, $E(y_{-1})$, . . . $E(y_{-m+1})$. If initial values for ε_o, . . . ε_{-q+1} are going to be specified a priori, the value zero is a natural choice since $E(\varepsilon_t) = 0$. This method of handling the presample values of the error term in the estimation of parameters in time series models is quite common (Box and Jenkins, Nelson).[9/]

It is important to recognize that $E(y_{t-j})$ in (19) is the unconditional expectation with respect to the dependence structure of the error term. After the autoregressive transformation to get (18) from the serial correlation specification of (16) and (17), we could define $E(y_t)$ as either conditional or unconditional with respect to y_{t-1} and y_{t-2} which appear on the right hand side of (18). We choose the unconditional specification for $E(y_t)$ which ignores the information in the data associated with the serially correlated disturbance, i.e., $E(y_t)$ is taken directly from (16) using $E(u_t) = 0$

instead of from (18) with y_{t-1} and y_{t-2} taken as known data. The advantage of this somewhat arbitrary choice is explained in the next section.

VI. Interpretation and Specification of the Disturbance Term

The simplified model of (12) with a first order serially correlated disturbance term is used to interpret the nonstochastic difference equation model. Since the disturbance term is

(22) $u_t = \rho u_{t-1} + \varepsilon_t$

we can write (12) as

(23) $y_t = \alpha + \beta\rho_{t-1} + \lambda E(y_{t-1}) + \rho u_{t-1} + \varepsilon_t$,

and ε_t obeys the classic assumptions. Now we know from the derivation of (10) and (11) that $E(y_t)$ is a function of lagged values of the exogenous variables, from t back to the first time period included in the sample, and $E(y_o)$. Therefore, $E(y_t)$ is strictly exogenous and not affected by the stochastic part of the model, i.e., not affected by lagged values of the dependent variable.

However, information about the current level of the dependent variable contained in the stochastic part of lagged values of this variable is carried by u_{t-1}, the lagged disturbance term. We think of this part of the model, lagged values of the disturbance term, as the endogenous component of the model. Specification of $E(y_t)$ as an unconditional expectation with respect to lagged values of the dependent variable, or equivalently lagged values of the disturbance term, permits a partitioning of the model into exogenous/endogenous components.

Let us rewrite (23) as

(24) $E(y_t) + u_t = \alpha + \beta p_{t-1} + \lambda E(y_{t-1}) + \rho u_{t-1} + \varepsilon_t$.

This equation is comprised of two components (1) a nonstochastic difference equation in the mean of y, and (2) a stochastic difference equation in the disturbance term of y. That is to say, (23) is simply the sum of (22) and (9), where respective sides of the equations are added together. This interpretation generalizes to higher order difference equations in an obvious way.

In the context of agricultural supply estimation, the difference equation in the residuals permits the model to reflect information contained in historic levels of output which cannot be captured by available data on exogenous variables. It is the author's opinion that this stochastic component will usually be of substantial importance in applications because of data limitations, minor specification errors such as nonlinearities in variables, and a host of rigidities and constraints in farmer's response which defy direct specification. A plausible rule of thumb would be to specify the order of serial correlation at least equal to the order of the nonstochastic difference equation.

Let us replace (22) by

(25) $u_t = -\theta\varepsilon_{t-1} + \varepsilon_t$

as the specification on the disturbance term, i.e., a first order moving average process. It is seen that the relationship in the stochastic part of the equation is no longer a difference equation, only one component of the error term in t-1 enters on the right hand side of (25). Since the $\{\varepsilon_t\}$ are assumed to be independently and identically distributed, the correlation between u_t and earlier values of the disturbance term ends with u_{t-1} and does not extend backwards in time at an exponentially declining magnitude as in the serial correlation model. In a supply response model, (25) implies that only the disturbance last year has an impact on this year's output and the disturbance in earlier years has no impact whatosever, not even indirectly. The assumption on the dependence structure is less dynamic, in a sense, than is the case with serial correlation.

We might expect the first order moving average specification on the disturbance term to be appropriate for an annual crop which is competitive with only other annual crops, particularly on a regional as opposed to an aggregate basis. A first order difference equation in $E(y_t)$ would also be quite likely in such an application since changes in crops would often be relatively easy. The application to U.S. soybean acreage reported in the next section substantiates this conjecture.

The general problem of specification of the structure of the error term is much the same as that encountered in Box-Jenkins type time series analysis, but somewhat more complicated because of the additional parameters associated with the exogenous variable and the difference equation. Not only are the additional parameters present, but they enter nonlinearly through the nonstochastic difference equation. One practical approach is to specify a fairly high order of serial correlation in the disturbance term since an invertible moving average process can be approximated in this way, as well as a combined serial correlation/moving average process (Nelson 1973 and 1976). Our emphasis in econometrics is frequently on the structural equation of the mean of the dependent variable with the systematic part of the error term viewed as merely a means to better estimate that structure. Nevertheless, parsimony of parameter numbers is always advantageous and frequently improves prediction, so that some exploration of the error structure might be justified.

VII. Illustrative Empirical Results

The results reported below are of necessity quite brief because of space limitations. The focus is on the relative advantage of nonstochastic over stochastic difference equations and the importance of specifying a sufficiently high order of difference equation to capture the dynamic structure in supply response.

U.S. Wheat Acreage Response

Acreage response of U.S. wheat producers provides an excellent illustration of the role of difference equations in measuring dynamic behavior and irreversibilites in supply response.

Rapid structural changes caused by variations in the government programs made it necessary to limit the sample period to 1961 - 77. There were not enough degrees of freedom to estimate a nonstochastic difference equation which requires three additional data points for a third order equation compared to the ordinary stochastic difference equation. The extra data points are required to handle the presample values $E(y_o)$, $E(y_{-1})$, and $E(y_{-2})$, either as parameters estimated from the sample or observed values y_o, y_{-1}, and y_{-2} before the structural change. If y_o, y_{-1}, and y_{-2} are used as estimators of the unkown parameters and these values extend back into the time series where the structure has changes, these lagged values are likely to be poor estimators and distort all the other parameter estimates. When using an ordinary stochastic difference equation, dropping back into the series where the structure has changed is not nearly as serious because errors in y_o, y_{-1}, y_{-2} only affect the first three observations in the sample, but with the nonstochastic difference equation, the errors are carried forward to all observations in the sample.

The recent work of Garst and Miller was used as a starting point for the acreage equation. One necessary change was to remove the jointly dependent variables associated with acreage diversion and the set-aside programs. Garst and Miller entered the acres signed up for these two programs as independent variables in the acreage seeded equation, but we would expect all three of these acreage variables to be jointly dependent. Therefore, additional dummy variables were introduced to replace the acres in the diversion and set-aside programs.

The variables in the acreage equation are:

X_1 = domestic allotment (1000 acres)

X_2 = no allotment-dummy (takes the value 1 after 1971, 0.26 in 1971, and zero otherwise -- see Garst and Miller)

X_3 = relaxed-allotment-dummy (takes the value 1 during 1965-70, 0.74 in 1971, and zero otherwise -- see Garst and Miller)

X_4 = diversion-dummy (takes the value 1 during 1963-66 and 1969-70, otherwise zero)

FGP = deflated feed grain price index (1967 base)

SP = deflated September price (1967 base)

P = deflated season average price (1967 base)

A = seeded acreage (1000 acres)

The deflator used is the annual index of prices paid by farmers for production items. When deflating price variables for a crop year, the calendar year index is a lagged deflator by a few months.

The season average wheat price is average price received by farmers including government subsidies. September price is the market price plus the average government subsidy payment per bushel for the crop year in which

September falls, which assumes farmers were able to anticipate the subsidy forthcoming for their crop.

A static model which seemed to give the best all-round performance was

$$
\begin{aligned}
A_t = & -52{,}256 + \underset{(5.5)}{.4365}\, X_{1t} + \underset{(6.3)}{30{,}648}\, X_{2t} - \underset{(0.1)}{94}\, X_{3t} \\
(26) \quad & -\underset{(4.2)}{4{,}352}\, X_{4t} + \underset{(1.2)}{5{,}807}\, FGP_t + \underset{(3.9)}{7{,}462}\, SP_{t-1} + \underset{(7.9)}{10{,}747}\, P_{t-1} \\
& + \underset{(8.9)}{11{,}945}\, P_{t-2} + \underset{(7.8)}{10{,}652}\, P_{t-3}
\end{aligned}
$$

$\bar{R}^2 = .985$, Std. error of estimate = 1190, and 7 degrees of freedom. The numbers in parentheses are absolute values of the t-ratios for the parmeter estimates above them. No serial correlation of any significance was found in the disturbance term.

The model would appear quite satisfactory except for two aspects, (1) the parameter estimates fluctuate a great deal as the sample period is changed by just an observation or two, and (2) prediction one year beyond the sample period is very poor in the latter part of the series when prices started to fall. The second deficiency is apparently a manifestation of the irreversibility phenomenon in supply response and the first problem of parameter instability is a symptom of specification error.

Difficulty was experienced in "discovery" of a dynamic model because a first or second order difference equation showed no improvement over the model in (26), but a third order difference equation showed a profound improvement. The fitted equation with a first order serial correlation specification on the disturbance term is

$$
\begin{aligned}
A_t = & -60{,}461 + \underset{(29.8)}{.5970}\, X_{1t} + \underset{(33.9)}{37{,}761}\, X_{2t} + \underset{(2.1)}{272}\, X_{3t} - \underset{(5.3)}{1353}\, X_{4t} + \underset{(8.3)}{13{,}144}\, FGP_t \\
(27) \quad & + \underset{(7.7)}{6047}\, SP_{t-1} + \underset{(2.4)}{691}\, P_{t-1} + \underset{(28.7)}{15{,}190}\, P_{t-2} + \underset{(21.5)}{.4502}\, A_{t-1} - \underset{(21.2)}{.5114}\, A_{t-2} \\
& + \underset{(26.3)}{.5619}\, A_{t-3} ,
\end{aligned}
$$

$\overline{R}^2$ = .9996, Std. error of estimate = 191, 3 degrees of freedom, and the point estimate of the serial correlation parameter is -0.924. The numbers in parentheses are approximate t-ratios for the respective parameters, approximate because the serial correlation parameter was estimated by nonlinear least squares simultaneously with the linear regression coefficients.

The dynamic model of (27) predicted with excellent accuracy as is illustrated in Table 1, and in addition, the parameter estimates changed very little as the sample was shortened, even down to zero degrees of freedom. For comparison, one-year-ahead predictions from the relatively static model of (26) gave errors of -271, 6227, and 8244 thousand acres in 1975, 1976, and 1977, respectively. The largest error from the dynamic model of (27) was 2560 thousand acres in 1977. The deflated September price associated with the 1977 prediction was below any experienced prices in the sample used to estimate the models, which makes the prediction unusually difficult in any case.

Elasticities of acreage response for various lengths of run are given in Table 2 for both (26) and (27) with price and acreage at their sample means. September price was treated as if it were current season average price in the calculation of these elasticities. Note that the long-run and one-year response elasticities are nearly the same for the two models, but the intermediate elasticities diverge considerably, especially the two-year response.

Table 1

Wheat Acreage Prediction from Third Order Difference Equation

Harvest Year	Seeded Acreage	1961-74 Sample Predicted Acreage	1961-74 Sample Error	1961-75 Sample Predicted Acreage	1961-75 Sample Error	1961-76 Sample Predicted Acreage	1961-76 Sample Error
		(units in thousand acres)					
1975	75,095	74,827	268				
1976	80,239	81,306	-1067	81,592	-1353		
1977	74,800	74,520	280	74,403	397	72,241	2560

Table 2

Acreage Response Elasticities for U.S. Wheat*

Years to respond	1	2	3	4	. . .	limit
Elasticities						
Static model	.25	.61	1.01	1.37		1.37
Dynamic model	.20	.32	.77	1.03		1.47

* Acreage and price at their means during 1961-77.

It is conjectured that the substantial negative serial correlation in the disturbance term for the dynamic model of (27) emanates from the cultural practice of summer-fallow in The Great Plains. An "unusually" large acreage in the aggregate this year implies an "unusually" small acreage next year because of the need to devote land to summer-fallow, ceteris paribus. This influence of summer-fallow on the disturbance term would be particularly important in the relatively wetter areas of The Great Plains where summer-fallow is a marginal practice. It would appear that the static model of (26) is too crude to pick up this behavior of the disturbance term; too many other confounding influences which are largely the result of specification errors must be relegated to the disturbance term.

U.S. Soybean Acreage Response

Soybean acreage response of U.S. farmers is used to make a comparison of nonstochastic and stochastic difference equations. The specification of the response equation is a generalization of that presented in (Houck, et.al.). The main change is that prices are deflated with the index of prices paid for production items, and the market price received for soybenas is included as a separate variable in addition to the ratio of soybean to corn prices used in (Houck, et.al.). Deflation of prices introduces a measure of the absolute profitability of growing field crops. The resulting model contains a variable for each soybean and corn prices together with the ratio of these two prices which serves as an interaction term to measure nonadditivity in the net price effects. The separate price for corn is "effective price support rate" as defined in (Houck, et.al.) for the years before 1971 and is market price lagged one year after 1971. Both soybean and corn market prices were lagged one year. As in (Houck, et.al.), soybean support price and effective diversion payment rate for corn in the current crop year were also used as independent variables.

Only summary measures of fit and statistical precision are given to contrast the nonstochastic and stocahstic difference equation models. The sample period used was 1952-77 and the difference equations are first order. Up to third order equations were tried with essentially no improvement in the fit or precision on the key economic variables. Results are summarized in Table 3. The average t-ratios are over the three variables: (1) lagged soybean price, (2) lagged ratio of soybean price to corn price, and (3) "effective price support rate" for corn (lagged corn price after 1971).

These results show the clear superiority of the nonstochastic difference equation for all three disturbance term specifications. It is noted that the t-ratios are only approximately distributed as the t-distribution because of nonlinearities in the parameters. All respective parameter estimates in the nonstochastic difference equations are essentially the same between the second order serial correlation and first order moving average error specification, but the precision is quite a little better in the latter. The only equation which gave substantially different point estimates of the parameters was the stochastic difference equation under the classical specifications.[10/]

Adequacy of a first order difference equation suggests that irreversibilities in soybean acreage response are relatively simple compared to wheat which required a third order equation. Apparently the mathematical expectation of acreage last year subsumes all the information in the systematic part of the acreage equation which is useful for explaining expected acreage this

Table 3

Summary Measures of Fit and Statistical Precision for U.S. Soybean Acreage Models

Model Description	$\bar{R}^2$	Standard Error of the Estimate (1000 acres)	Average t-ratio on Price Variables a/
Classical Error Term			
Nonstochastic Dif. Eq.	.9931	1120	3.84
Stochastic Dif. Eq.	.9862	1578	2.82
Second Order Serial Correlation			
Nonstochastic Dif. Eq.	.9944	1005	6.43
Stochastic Dif. Eq.	.9900	1346	3.85
First Order Moving Average Error b/			
Nonstochastic Dif. Eq.	.9968	766	7.42
Stochastic Dif. Eq.	.9934	1095	3.95

a/ Lagged prices of soybeans, corn, and the ratio of soybean to corn price.

b/ The moving average error parameter had to be constrained to acheive invertibility and was set equal to 0.95 in both equations.

year. Recent changes in acreage, or the rate of change, are relatively unimportant.

The first order moving average error term also suggest a relatively simple dynamic structure in the stochastic part of the acreage equation since the disturbance in year t is correlated with the disturbance only back to year t - 1. Since the dependence structure of the disturbance term in a crop acreage response equation is probably dominated by aggregate crop rotation constraints, these results suggest a rapid adjustment process in the rotations containing soybeans. This should be no surprise for readers familiar with farm level production of soybeans. In contrast, the role of summer-fallow in wheat rotations in The Great Plains would suggest that fluctuations in the disturbance term would require several years for dissipation, as characterized by first order serial correlation.

Short-run price elasticity of acreage response at mean price and acreage during 1952-77 was estimated at 0.53 from the nonstochastic difference equation with moving everage error; the long-run elasticity implied by the model was 4.43. The short and long run cross elasticities with respect to lagged corn price were -0.53 and -4.46, respectively, essentially the same as own price elasticity except opposite in sign. The partial derivative of the acreage response equation with respect to soybean price is essentially the same as that obtained by Kenyon and Evans (identically the same for two significant digits); their reported short-run elasticity is higher because of the mean price and acreage differences used for the calculations.

U.S. Wheat Yield Response

Aggregate yield per seeded acre was used as the dependent variable and the independent variables were linear trend, average price received by farmers per bushel of wheat (including subsidies), and acreage seeded. A sixth order lag was used on wheat price, starting with the current year's price, and the prices for years t - 2 and t - 4 were deleted because of confounding of the estimators with the other lagged prices.[11/] The difference equation and serial correlation in the disturbance were each specified as third order, making a total of 13 independent parameters to estimate. The sample period used after accounting for the autoregressive transformation to estimate the serial correlation parameters was 1952-77, leaving 13 degrees of freedom.

Results contrasting the nonstochastic and stochastic difference equations are presented in Table 4 for a classical disturbance term as well as third order serial correlation. The advantage of the nonstochastic difference equation is rather profound in this application because of the inherent variablity of crop yeilds; forcing the lagged disturbance term to have the same net effect on yield as the lagged expectation of yield poses a serious constraint, and this is exactly what the stochastic difference equation specification does. On the other hand, the third order nonstochastic difference equation, jointly with the same order of serial correlation specified on the disturbance term, provides a model which involves a pair of separate difference equations, one each for the systematic part of the yield equation and the random disturbance term.

Montana Beef Cattle Breeding Stock

Some very preliminary results from research by Randy Rucker is reported to show the apparent potential of nonstochastic difference equations in explaining changes in beef breeding herd inventories on January 1 each year. The model for Montana used a second order lag on annual hay production and average price received for calves in Montana during the last quarter of the

Table 4

Summary Measures of Fit and Statistical Precision for U.S. Wheat Yield Models

Model Description	$\bar{R}^2$	Standard Error of the Estimate (bushels/acre)	Average t-ratio on Price Variables
Classical Error Term			
Nonstochastic Dif. Eq.	.9039	1.63	1.93
Stochastic Dif. Eq.	.8662	1.92	0.82
Third Order Serial Correlation			
Nonstochastic Dif. Eq.	.9631	0.87	7.95
Stochastic Dif. Eq.	.8891	1.50	2.62

the calendar year; the other explanatory variable is the ratio of choice cattle price to corn price in Omaha. Second order difference equations with the same order of serial correlation in the distrubance term appeared adequate.

Results are for the sample period 1951-78. The respective measures for the nonstochastic and stochastic difference equations are as follows: adjusted R-squared, .9967 and.9899; standard error of the estimate (1000 head), 19.4 and 33.9; average t-ratio on prices, 13.5 and 2.82.

III. Summary and Conclusions

The complex dynamic structure of agricultural supply response is best approximated by difference equations in the response variables jointly with relatively low order finite distributed lags on the exogenous variables. A distinction is made between stochastic and nonstochastic difference equations, and it is argued that the nonstochastic version is better adapted to measurement of supply response.

However, the ulitimate dynamic model for supply estimation uses a combination of stochastic and nonstochastic difference equations by a partitioning of response into exogenous and endogenous components. The exogenous component is a nonstochastic difference equation in a variable defined as the unconditional mathematical expectation of the response variable, while the endogenous component is a stochastic difference equation in the random disturbance term. The latter component is introduced in practice by specifying the disturbance term with a serial correlation structure, or possibly, a moving average error structure which does not exactly yield a stochastic difference equation for the disturbance term.

We speak of the dependence structure (serial correlation or moving average error) in the disturbance term as an endogenous component because its information content emananates from lagged values of the disturbance term, i.e., lagged values of the stochastic component of the supply response variable. Consequently, the informational value of the dependence structure of the disturbance term is generated internally from lagged values of the dependent variable, and is conceptually endogenous in nature within the constructs of the model.

Statistical estimation of parameters in these dynamic models is feasible and not particularly expensive or burdensome on the analyst. Nonlinear least squares algorithms handle the problem quite easily by recursively calculating partial derivatives of the response variable with respect to the parameters of the equation[12/]; the Marquardt algorithm is recommended particularly with the improvements developed in Fletcher. Starting values for the parameters are easily obtained by replacing lagged expectations of the dependent variable by observed lagged values. Under the assumption of normality on the disturbance term the nonlinear least squares estimates are maximum likelihood.

So-called problems of irreversibility in supply response are obviated by specification of a sufficiently high order difference equation. When symptoms of irreversibility are found in an estimated regression equation, it is merely a manifestation of specification error associated with the dynamic structure of the response equation.

Summary results are presented for applications to U.S. wheat and soybean acreages, U.S. wheat yield, and Montana beef breeding herd inventories. These results verify the logical developments of the paper and show the order of magnitude of improvement that can be expected in empirical research. Apparent improvements in statistical precision are substantial and would suggest a much improved methodology for statistical estimation of supply response from time series data.

The nonstochastic difference equation models would appear to be especially vulnerable to confounding influences of seasonal variation in the data. For example, a second order nonstochastic difference equation with complex roots can be solved explicitly in time as a sinusoidal function of the general form $A \cos(vt + \theta)$. Such a relationship could easily track the periodic behavior of seasonal variation instead of measuring the general dynamic structure of supply response. Therefore, special care must be exercised to remove the influence of seasonal variation from the data before building a dynamic supply model (see Sims 1974a).

FOOTNOTES

1/ The reader is referred to Griliches' 1967 survey article or (Dhrymes, 1971a) for an analysis of various lag distributions.

2/ We do not consider problems of jointly dependent variables and simultaneous equation systems.

3/ If we were to add a disturbance term to (1), making x_t^* stochastic, then the expected values in (3) would have to be interpreted as conditional, given x_t^*. In this case, the disturbance term u_t in (5) would be comprised of two components, one emanating from a disturbance added to (1) and the second from additional variation in x_t caused by such things as weather, minor omitted variables, etc. In either specification, the disturbance term in (5) should have essentially the same properties, and the choice between the two specifications is purely arbitrary.

4/ In order for this conclusion to hold, we must be careful to define $E(x_t)$ as the conditional expectation, given x_{t-1} in (5)′ or given u_{t-1} in (6). This point will be clarified in section VI.

5/ Kmenta obtains (9) from (7), where u_t is specified as $u_t = \varepsilon_t - \lambda\varepsilon_{t-1}$ which emanates from the Koyck transformation, but he seems to use (9) only as a pedagogical device to derive (11), see (Kmenta, p.481)

6/ Some other ideas on ways to deal with the presample data problem and to estimate $E(y_o)$ are explored by Pesaran.

7/ A disadvantage is that stability constraints on the difference equation are difficult to impose as compared to the direct search procedure. One should always check for stability after getting the least squares estimates, and in the case of serial correlation models which follow, the resulting parameters should be analyzed to check for a stationary stochastic process. The moving average error models discussed later need to be checked for invertibility (Nelson 1973).

8/ The two observations lost by the autoregressive transformation can be salvaged by application of a special transformation to each the first and second data points (Schmidt). Maximum likelihood estimates for this larger sample will not be the same as those obtained by minimizing the error sum of squares in (18), but will approach the least squares estimates for large samples. The difference between least squares and maximum likelihood estimates would appear to be trivial for small samples unless the boundary for stability of the serial correlation process is approached. One problem with saving the observations lost by the autoregressive transformation is that an assumption must be made that the stochastic process generating the residuals u_t has been in operation for an extended period of time and is stationary (Theil, footnote p.253). In agricultural supply analysis, we cut off the sample at some point because we suspect a serious change in structure, consequently, it would seem rather dubious to try to salvage a couple of observations in light of the tacit assumption required to justify doing so.

9/ The computer program developed by the author with the able assitance of Stuart Townsend sets the presample values of the $\{\varepsilon_t\}$ equal to zero and those for the $\{E(y_t)\}$ equal to y_t. This method saves some degrees of freedom in small samples and the results are asymptotically independent of the method used to deal with the presample values. A problem exists in using least squares estimation for the moving average error model because invertibility is not imposed in the estimation precedure and the method only approaches maximum likelihood for large samples (Wallis). Therefore, the least squares method can encounter problems when the parameters are close to the invertibility boundary and sample size is small to moderately large.

10/ Problems associated with lagged values of the dependent variable as explanatory variables in regression when there is serial correlation in the residuals are well known (Fuller and Martin). Such models are also likely to be misleading in that there is a tendency for specification errors to be hidden (Sims 1974b, p. 300).

11/ Additional research is underway to try to find a more acceptable constraint on the distributed lag parameters for price. These parameter estimators are highly correlated in a systematic way which suggests a confounding caused by the role of summer-fallow in the crop rotations of The Great Plains.

12/ This same approach to parameter estimation could be used when producer expected price is defined as a known function (except for unknown parameters) of lagged prices. Expected price is simply treated as an unobservable variable comparable to the lagged expectation of the dependent variable in the nonstochastic difference equation. This approach would be particularly useful in the risk models for supply modelling which Richard Just has developed (Just 1974a, 1974b, and 1977).

References

Askari, Hossein and John Thomas Cummings, "Estimating Agricultural Supply Response with the Nerlove Model: A Survey," *International Economic Review*, 18(1977): 257-92.

Baritelle, John L. and David W. Price, "Supply Response and Marketing Strategies for Deciduous Crops," *Amer. J. Agr. Econ.* 56(1974): 245-53.

Behrman, Jere R., "Price Elasticity of the Marketed Surplus," *J. Farm Econ.*, 48(1966): 875-893.

Box, George E.P. and Gwilym M. Jenkins, *Time Series Analysis: Forecasting and Control*, Rev. Ed., San Francisco: Holden - Day, 1976.

Cassels, John M. "The Nature of Statistical Supply Curves," *Journal of Farm Economics*, 15(1933): 378-87.

Chen, Dean, Richard Courtney, and Andrew Schmitz, "A Polynomial Lag Formulation of Milk Production Response," *Amer. J. Agr. Econ.*, 54(1972): 77-83.

Dhrymes, Phoebus, J., *Distributed Lags: Problems of Estimation and Formulation*, San Francisco: Holden - Day, 1971a.

Dhrymes, P.J., "On the Strong Consistency of Estimators for Certain Distributed Lag Models with Autocorrelated Errors," *International Economic Review*, 12(1971b): 329-342.

Draper, N.R. and H. Smith, *Applied Regression Analysis*, New York: Wiley, 1966.

Fletcher, R., *A Modified Marquardt Subroutine for Non-Linear Least Squares*, United Kingdom Atomic Energy Authority, Research Group Report AERE - R6799, Harwell, Berkshire, 1971. (Available from H.M. Stationery Office)

Garst, Gail D. and Thomas A. Miller, "Impact of the Set-Aside Program on the U.S. Wheat Acreages," *Agricultural Economics Research*, 27(1975): 30-37.

Fuller, Wayne A. and James E. Martin, "The Effects of Autrocorrelated Errors on the Statistical Estimation of Distributed Lag Models," *J. Farm Econ.*, 43(1961): 71-82.

Grether, David M., "A Note on Distributed Lags, Prediction, and Signal Extraction," *Econometrica*, 45(1977): 1729-34.

Griliches, Zvi, "Estimates of the Aggregate U.S. Farm Supply Function," *J. Farm Econ.*, 42(1960): 282-93.

Griliches, Zvi, "Distributed Lags: A Survey," *Econometrica*, 35(1967): 16-49.

Houck, James P., Martin E. Abel, Mary E. Ryan, Paul W. Gallagher, Robert G. Hoffman, and J.B. Penn, *Analyzing the Impact of Government Programs on Crop Acreage*. USDA Technical Bulletin No. 1548, Washington D.C., 1976.

Houck, James P., "An Approach to Specifying and Estimating Nonreversible Functions," Amer. J. Agr. Econ., 59(1977): 570-72.

Jorgenson, Dale, "Rational Distributed Lag Functions," Econometrica, 34(1966): 135-49.

Just, Richard E., Econometric Analysis of Production Decisions with Government Intervention: The Case of California Field Crops, Giannini Foundation Monograph 33, University of California Berkeley, 1974a.

Just, R.E., "An Investigation of the Importance of Risk in Farmer's Decisions," Amer. J. Agr. Econ., 46(1974b) 14-25.

Just, Richard E., "Estimation of an Adaptive Expectations Model," International Economic Review, 18(1977): 629-644.

Kmenta, Jan, Elements of Econometrics, New York: Macmillan Co., 1971.

Kenyon, David E. and R.S. Evans, Short-Term Soybean Acreage Projection Model Including Price and Policy Impacts, Research Division Bulletin 106, Virginia Polytechnic Institute & State University, Blacksburg, Virginia, 1975.

Klein, L.R., "The Estimation of Distributed Lags," Econometrica, 26(1958): 553-65.

Kulshreshtha, Surendra N., "An Analysis of the Canadian Cattle Supply Using Polynomial Distributed Lags," Canadian J. Agr. Econ., 24(1976): 1-14.

Nelson, Charles R., Applied Time Series Analysis, San Francisco: Holden - Day, 1973.

Nelson, Charles R., "The Interpretation of R^2 in Autoregressive - Moving Average Time Series Models," American Statistician, 30(1976): 175-180.

Nerlove, Marc, "Estimates of the Elasticities of Supply of Selected Agricultural Commodities," J. Farm Econ., 38(1956): 496-509.

Nerlove, Marc, The Dynamics of Supply: Estimation of Farmers' Response to Price, Baltimore: Johns Hopkins Press, 1958.

Pesaran, M.H., "The Small Sample Problem of Truncation Remainders in the Estimation of Distributed Lag Models with Autocorrelated Errors," International Economic Review, 14(1973): 120-31.

Russell, Bertrand, Human Knowledge Its Scope and Limits, London: George Allen and Unwin Ltd., 1948.

Schmidt, Peter, "Estimation of a Distributed Lag Model with Second Order Autoregressive Disturbances: A Monte Carlo Experiment," International Economic Review, 12(1971): 372-380.

Sims, Christopher, "Seasonality in Regression," Journal American Statistical Association, 69(1974a): 618-26.

Sims, Christopher, "Distributed Lags," in Frontiers of Quantitative Economics, Vol. II, ed. M.D. Intrilligator and D.A. Kendrick, North - Holland, 1974b, pp. 289-338.

Solow, R.M., "On a Family of Lag Distributions," Econometrica, 28(1960): 393-406.

Theil, Henri, Principles of Econometrics, New York: John Wiley and Sons, 1971.

Traill, Bruce, David Coleman, and Trevor Young, "Estimating Irreversible Supply Functions," Amer. J. Agr. Econ., 60(1978): 528-31.

Wallis, Kenneth F., "Multiple Time Series Analysis and the Final Form of Econometric Models," Econometrica, 45(1977): 1481-97.

Wolffram, R., "Positivistic Measures of Aggregate Supply Elasticities: Some New Approaches -- Some Critical Notes," Amer. J. Agr. Econ., 53(1971): 356-59.