89-16

UCD Department of Agricultural Economics

Supply and demand -- Mathematical models

WORKING PAPER SERIES

#7135

University of California, Davis
Department of Agricultural Economics

IRREVERSIBLE SUPPLY FUNCTIONS: CONCEPTS AND ESTIMATION

Oscar R. Burt

November 10, 1989

Working Paper No. 89-16

Irreversible Supply Functions: Concepts and Estimation

by

Oscar R. Burt
University of California, Davis
Department of Agricultural Economics

Irreversible Supply Functions: Concepts and Estimation

Fixed asset theory in agricultural production has a long tradition in our profession (Johnson), and has led to the term irreversible supply functions. During the last decade or so there has been a growing interest in dynamic demand equations which also might exhibit irreversibilities. The primary means for statistically testing and modeling irreversible functions has been that proposed by Wolffram and its modification by Houck. Ironically, this method evolved for purposes of supply response estimation (Tweeten and Quance), but appears to be appropriate for demand and not supply equations.[1]

A fundamental problem with Wolffram's model for supply response is that changes in prices, in and of themselves, cause a permanent change in output. A sequence of price movements from an initial level with a return to that level, followed by no further price changes, will not ultimately lead to the same output as would have occurred if price had remained constant at the initial value. There is no long-run equilibrium associated with a given price level.[2] Fixed asset theory in supply response would imply only short-run asymmetries to account for depreciation of capital stocks. Of course, technological change will permanently shift the long-run supply function, but this is a separate matter from price changes per se doing so.

On the other hand, the habit formation hypothesis in demand theory would be compatible with Wolffram's model because a change in consumption associated with price changes could precipitate a change in tastes. Supposedly the consumer's history of consumption affects his tastes and preferences, thus the term habit formation. A period of relatively low prices for a good could stimulate a permanent increase in the quantity desired at a given price, and vice versa. But the author sees no particular reason to expect asymmetry with respect to positive and negative changes in prices in general as in supply response. The case of drugs and narcotics is one small set of goods where irreversibility probably prevails.[3] The stock-adjustment phenomenon in consumer behavior with respect to durable goods is essentially the same as for capital stocks in supply response. This paper focuses on the dynamics of supply response, but is applicable to the stock-adjustment case of dynamic demand equations.

The next section reviews the Wolffram model of irreversible functions and interprets this model and Houck's modification in the context of how initial conditions are estimated for the dynamic process implied by the regression equations. The third section presents a distributed lag model for asymmetric responses which does have a well defined equilibrium for a given price level. Aggregate crop output in the United States is analyzed with this model and its symmetric counterpart for comparison. The fourth section is a general discussion of dynamic supply response in which it is argued that most apparent irreversibilities are a reflection of inadequate modeling of capital stocks and other latent state variables which underlie the production process. The paper closes with some concluding remarks on econometric specification and estimation of supply response relationships.

Wolffram's and Houck's Methods

A basic weakness in Wolffram's argument to support his method, also in Houck's reformulation of it, was absence of a disturbance term in the equation used for analysis. Suppose the time series sample is comprised of n observations and the specification desired is such that increases and decreases in the explanatory variable x have different slopes in a linear relationship with the dependent variable y, and z denotes a second explanatory variable which enters the equation with an ordinary linear response. Following Houck except for addition of a disturbance term, u, the differenced form of such a regression equation can be written,

$$(1) \qquad \Delta y_t = \beta_1 \Delta^+ x_t + \beta_2 \Delta^- x_t + \gamma \Delta z_t + \Delta u_t , \qquad t = 2, 3, \ldots n,$$

where Δ is the difference operator such that $\Delta y_t = y_t - y_{t-1}$, and $\Delta^+ x_t$ and $\Delta^- x_t$ denote Δx_t when it is positive and negative, respectively, and zero otherwise; β_1, β_2, and γ are unknown parameters. If Δu_t had properties which made statistical estimation easy, there would be little more to discuss; (1) would be estimated as written.

But Δu_t is likely to have a first order moving average component, and certainly will have if u_t has the classic properties, or is autoregressive without a unit root. Therefore, interest lies in the integrated equivalent of (1) obtained by summing both sides from i=2 to n to get

$$(2) \qquad y_t - y_1 = \beta_1 \left(\sum_{i=2}^{t} \Delta^+ x_i \right) + \beta_2 \left(\sum_{i=2}^{t} \Delta^- x_t \right) + \gamma (z_t - z_1) + (u_t - u_1).$$

Transposing y_1 to the right hand side and collecting terms gives

(3) $y_t = (y_1 - u_1) + \beta_1 X_t^+ + \beta_2 X_t^- + \gamma(z_t - z_1) + u_t,$

where X_t^+ and X_t^- denote the cumulative sums of Δ^+x and Δ^-x in (2).

If u_t has the classic properties, an obvious way to estimate (3) is by linear least squares and the constant term is an estimate of the initial condition expression, $y_1 - u_1$. Houck's analysis without a disturbance term led him to use y_1 as an a priori initial condition; then a disturbance is added for estimation purposes. The intercept in (3) is $y_1 - u_1 = E(y_1)$, where $E(\cdot)$ is the expectation operator. Therefore, Houck's approach uses y_1 as an a priori estimator of $E(y_1)$ by forcing the intercept to zero and defining the dependent variable as $y_t - y_1$, $t = 2, 3, \ldots n$.

Apparently Houck thought his method would give the same point estimates of parameters as Wolffram's: "This method is consistent with the Wolffram technique but is operationally clearer." (Houck, p. 570). This is not the case except for the example constructed by Wolffram, or others like it which are deterministic. Wolffram's method saves the first observation and could be implemented by assigning zero to each variable on the right hand side of (3) for the first observation with the others unchanged, and an intercept is estimated. Thus Wolffram's method could be represented by

(4) $y_1 = \eta_1 + u_1$

(5) $y_t = \eta_1 + \beta_1 X_t^+ + \beta_2 X_t^- + \gamma(z_t - z_1) + u_t, \quad t = 2, 3, \ldots n,$

where the intercept η_1 is $y_1 - u_1 = E(y_1)$. In an example like Wolffram's where the variance of u_t is zero, the least squares estimate of η_1 is simply y_1 which shows why Houck's method would give the same results as Wolffram's for the latter's example.

As a practical matter, one would not expect much difference in the properties of these alternative estimators, but Wolffram's does use all the data to estimate η_1, which would seem to be an advantage. An unrepresentative observation on y_1 could produce poor estimates in a small sample using Houck's approach. Frequently the beginning of a time series sample is determined by assumptions about structural change and the risks of going back too far. Therefore, the first observation would tend to be an outlier more frequently than later ones.

If data are available on the independent variables for the presample observation $t = 0$, a third estimation method would be to use (5) for $t = 1, 2, \ldots n$, with $\eta_0 = y_0 - u_0$ replacing η_1 as the intercept. Note that y_0 is not involved in the estimation of parameters because (4) is not used in the estimation. As a general statement, it would be preferable to estimate η_0 jointly with the other parameters instead of setting it equal to y_0 using Houck's procedure, but there is not much to distinguish this third method from Wolffram's. Whether such considerations are of much importance in applications using small samples is an empirical question.

Serious interpretation problems exist for empirically fitted equations of the type given in (3) because the irreversible aspect is very likely to measure an autonomous trend unaccounted for by the independent variable set. Suppose z_t is a positively trended variable, or even a linear trend with positive slope, then the inequality $\beta_1 > \beta_2$ will allow the two signed variables in X to provide a rachet effect which mimics a positive trend. The end result is likely to be a special case of multicollinearity but more subtle than usual to detect. Frequently, the a priori basis for a trend variable is weak and it is used reluctantly to improve the specification, but in these circumstances, an estimated irreversible equation may appear to be appropriate when actually the response is symmetric and there is a missing variable from the equation. The asymmetric models analyzed in the next section are also vulnerable to this ambiguity of trends, but the problem would appear to be less acute because the specification forces a long run equilibrium response for a given level of price and this anchors the irreversible part.

Asymmetric Models with Equilibria

A serious deficiency noted for the Wolffram approach in modeling supply response is absence of an implied long run equilibrium for the dependent variable of the dynamic regression equation. This problem can be corrected by specifying a distributed lag on the signed price changes and introducing the level of price as a concommitant variable, where the dependent variable, quantity supplied, is also in levels. In general terms,

$$(6) \qquad q_t = \alpha + \sum_{j=1}^{\infty} \beta_j \Delta^+ p_{t-j} + \sum_{j=1}^{\infty} \gamma_j \Delta^- p_{t-j} + \delta p_{t-1} + u_t,$$

where $\Delta^+ p_{t-j}$ and $\Delta^- p_{t-j}$ are defined the same way as $\Delta^+ x_t$ and $\Delta^- x_t$ were in (1); while q and p are quantity and price, respectively. It is assumed that price last period is the nearest price in time to affect output.

If we were to replace the two signed differenced variables in price by simply Δp_{t-j} in a single summation, it can be shown that the resulting equation,

$$(7) \qquad q_t = \alpha + \sum_{j=1}^{\infty} \beta_j \Delta p_{t-j} + \delta p_{t-1} + u_t,$$

would be equivalent to the parameterization (see Burt 1989),

$$(8) \qquad q_t = \alpha + \sum_{j=1}^{\infty} \delta_j p_{t-j} + u_t,$$

i.e., a general distributed lag equation with symmetric response. The equilibrium for these three equations is obtained by taking price fixed at p_*, which implies $\Delta p_{t-j} = \Delta^+ p_{t-j} = \Delta^- p_{t-j} = 0$, $j = 1, 2, \ldots$. Clearly the equilibrium for (6) and (7) is $q_* = \alpha + \delta p_*$, in an expected value sense, and for (8) the coefficient on p_* is $\delta_1 + \delta_2 + \ldots$ which is assumed to be finite and must equal δ in (6) and (7).

With this background, it is seen that (6) provides a general framework to model asymmetric response such that a long-run equilibrium exists for the expected value of the dependent variable.[4] We can draw on all of the econometric literature for parsimonious parametrizations of the general linear distributed lag model in (8). One of the simplest is the geometric lag, which with extra terms added in the lagged independent variable, provides considerable flexibility. The symmetric model of (7) with a geometric lag imposed on the Δp_{t-j} is

$$(9) \qquad q_t = \alpha + \beta(\Delta p_{t-1} + \lambda \Delta p_{t-2} + \lambda^2 \Delta p_{t-3} + \ldots) + \delta p_{t-1} + u_t,$$

which can be shown to be equivalent to the parameterization

$$(10) \qquad q_t = \alpha + \gamma_1 p_{t-1} + \gamma_2 (p_{t-2} + \lambda p_{t-3} + \lambda^2 p_{t-4} + \ldots) + u_t,$$

where $\gamma_1 = \beta + \delta$ and $\gamma_2 = -\beta(1 - \lambda)$.

Clearly (10) allows a free parameter on p_{t-1} and a geometrically constrained lag on higher order lagged prices. Putting the geometric lag on Δp_{t-j} in (9) and including the term with p_{t-1} is therefore seen to give a specification with more flexibility than a regular geometric lag.

The asymmetric analogue of (9) or (10) is

$$(11)\quad q_t = \beta_0 + \beta_1\left[\Delta^+ p_{t-1} + \lambda_1\Delta^+ p_{t-2} + \lambda_1^2\Delta^+ p_{t-3} + \ldots\right] + \beta_2\left[\Delta^- p_{t-1} + \lambda_2\Delta^- p_{t-2} + \lambda_2^2\Delta^- p_{t-3}\cdot\cdot\cdot\right] + \beta_3 p_{t-1} + u_t,$$

where unknown parameters are β_0, β_1, β_2, β_3, λ_1, λ_2, and λ_1 and λ_2 are assumed to be less than one in absolute value, and most likely positive. This model imposes a separate geometric lag on each of the signed price change variables.

If price is held constant for an extended period, the terms in square brackets approach zero. Therefore, equilibrium quantity for a given price at p_* is

$$(12)\quad q_* = \beta_0 + \beta_3 p_*.$$

It should be clear from the details of the above model that a better term than irreversible response would be short-run asymmetries in response, because the temporal process is reversible asymptotically.

Expected restrictions on the parameters would be: $\beta_1 < 0$, $\beta_2 < 0$, $\beta_3 > 0$, $0 \le \lambda_1, \lambda_2 < 1$, and $|\beta_1| \le |\beta_2|$. The short-run marginal effects of p_{t-j} are obscured in (11) because p_{t-j} appears twice in the equation, with each a positive and negative sign. When there is no change in signs on Δp between terms t–(j–1) and t–j, it is readily seen that

(13) $\partial q_t/\partial p_{t-j} = -\beta_i(1-\lambda_i)\,\lambda_i^{j-2}, \quad j > 1,$

where i=1 if $\Delta p_{t-(j-1)}$ and Δp_{t-j} are both positive, and i=2 for the comparable terms in Δp when both are negative.

When there is a change in sign on $\Delta p_{t-(j-1)}$ and Δp_{t-j}, the results can be deduced from the sequence,

$$(14)\quad \beta_1 \lambda_1^{i-2} \left[p_{t-(i-1)} - p_{t-i}\right] + \beta_2\lambda_2^{i-1} \left[p_{t-i} - p_{t-(i+1)}\right] + \beta_1 \lambda_1^{i} \left[p_{t-(i+1)} - p_{t-(i+2)}\right],$$

where the signs of the bracketed terms are +, –, + in the order they appear. From (14), it is seen that when the sign change from period t–(j–1) to t–j is positive to negative and negative to positive, the partial derivatives for i > 1 are, respectively,

$$(15)\quad \partial q_t/\partial p_{t-j} = -\beta_1 \lambda_1^{j-2} + \beta_2 \lambda_2^{j-1} \quad \text{(sign + to –)}$$

$$(16)\quad \partial q_t/\partial p_{t-j} = -\beta_2 \lambda_2^{j-2} + \beta_1 \lambda_1^{j-1} \quad \text{(sign – to +).}$$

Since p_{t-1} occurs once and only once among the variables $\Delta^+ p_{t-j}$ and $\Delta^- p_{t-j}$, $j = 1, 2, \ldots,$

$$(17)\quad \partial q_t/\partial p_{t-1} = \begin{cases} \beta_1 + \beta_3 \text{ if } \Delta p_{t-1} > 0 \\ \beta_2 + \beta_3 \text{ if } \Delta p_{t-1} < 0 \end{cases}$$

Both β_1 and β_2 being negative makes the one period response to an increment in p_{t-1} less than the long-run effect, β_3. When the price changes between periods t–(j–1) and t–j are of the same sign, the marginal net effects of p_{t-j}

given by (13) are positive and smaller than the associated $|\beta_1|$ or $|\beta_2|$, for rising and falling prices, respectively, since $0 < \lambda_i < 1, i = 1, 2$. When Δp_k changes from positive to negative between periods t–(j–1) and t–j, the marginal net effect of p_{t-j} shown in (15) is negative because λ_1 is smaller than λ_2, $|\beta_1| \leq |\beta_2|$, and $\beta_i < 0, i = 1, 2$. But when Δp_k changes from negative to positive between periods t–(j–1) and t–j, the marginal net effect of p_{t-j} given in (16) is not necessarily positive because of the exponents on λ_1 and λ_2.

It is convenient to introduce the lag operator to discuss practical estimation procedures for (11). Let L be such an operator on the subscript of a time series variable x_t such that $L^j x_t = x_{t-j}$. If

(18) $$z_t = x_t + \lambda x_{t-1} + \lambda^2 x_{t-2} + \dots,$$

then lagging both sides of (18), multiplying by λ, and subtracting the results from the respective sides of (18) yields

$$z_t - \lambda z_{t-1} = (1-\lambda L)\, z_t = x_t$$

because all the terms on the right hand side cancel except for x_t. Therefore, z_t can be written as

(19) $$z_t = x_t/(1-\lambda L)$$

if $|\lambda| < 1$ so that the series in (18) converges. Using these results, (11) can be written as

(20) $$q_t = \beta_0 + \beta_1 \Delta^+ p_{t-1}/(1-\lambda_1 L) + \beta_2 \Delta^- p_{t-1}/(1-\lambda_2 L) + \beta_3 p_{t-1} + u_t.$$

Taking expectations of both sides of (20) and multiplying by $(1 - \lambda_1 L)(1 - \lambda_2 L)$ gives

$$(21)\quad (1-\lambda_1 L)(1-\lambda_2 L)E(q_t) = (1-\lambda_1 L)(1-\lambda_2 L)\beta_0 + \beta_1(1-\lambda_2 L)\Delta^+ p_{t-1} + \beta_2(1-\lambda_1 L)\Delta^- p_{t-1} + \beta_3(1-\lambda_1 L)(1-\lambda_2 L)p_{t-1}.$$

Adding the disturbance u_t to both sides and rearranging terms yields

$$(22)\quad q_t = (1-\lambda_1)(1-\lambda_2)\beta_0 + \beta_1(\Delta^+ p_{t-1} + \lambda_2\Delta^+ p_{t-2}) + \beta_2(\Delta^- p_{t-1} + \lambda_1\Delta^- p_{t-2}) + \beta_3\left[p_{t-1} - (\lambda_1 + \lambda_2)p_{t-2} + \lambda_1\lambda_2 p_{t-3}\right] + (\lambda_1 + \lambda_2)E(q_{t-1}) - \lambda_1\lambda_2 E(q_{t-2}) + u_t,$$

where we use $q_t = E(q_t) + u_t$. Although $E(q_{t-1})$ and $E(q_{t-2})$ are unobservable, they are implicitly defined as a function of the right hand side variables lagged back to the beginning of the sample, the unknown parameters in (22), and two initial condition parameters, $E(q_0)$ and $E(q_{-1})$. Details of estimation procedures for similar types of models are given in (Burt 1980), and in principle, this is a generalization of the transfer function model (Box and Jenkins; Harvey).

The above model was applied to the aggregate output index for all crops in the United States over the period 1914-1951. This historical period was selected because it preceded the large government intervention programs after World War II which often included acreage controls on various crops jointly with subsidies and market price distortions. The symmetric model specification is almost the same as used in La France and Burt which is a modification of that first used by Griliches. The one exception is an extra free parameter required to apply the geometric lag to Δp_{t-j} in (9). The price variable is the ratio of the indices of prices received to prices paid for crops, and the equation also contains weather index and trend variables which enter without any distributed lag response, i.e., they are just added terms in (20) like the level of price, p_{t-1}.

Results for several estimated equations are reported in Table 1. The first equation is the symmetric model for comparison. In the second equation, the signs and relative ordering for λ_1 and λ_2 are met, but $|\hat{\beta}_1| > |\hat{\beta}_2|$ which implies a greater first year response to a decrease in price than an increase, i.e., $\hat{\beta}_3 + \hat{\beta}_1 = (.241 - .196) < (.241 - .155) = \hat{\beta}_3 + \hat{\beta}_2$. The third equation has $\hat{\beta}_1$ constrained to equal $\hat{\beta}_2$ so that first year response is the same for price increases and decreases. This would seem quite plausible in that short-run response is effected primarily through adjustment of variable inputs. The prices used here are for March in the calendar year of production. Since the adjusted R-squared is smaller in the second and third equations than in the first, we know that the extra parameters which allow for asymmetric response are insignificant at the usual test levels.

The qualitative structure in the third equation is quite plausible. This can be seen by reviewing the implied lagged response structure of equations one and three. We convert the coefficients to weights which sum to one for easy comparison. For the symmetric model, the first five coefficients on lagged price (p_{t-j}, j=1, 2, ... 5) are: .295, .163, .125, .096, .074. For the asymmetric model, the same coefficients for rising and falling prices are respectively: .258, .169, .131, .101, .078 and .258, .059, .054, .050, .046. It is seen that the lag structure for rising prices is close to that for the symmetric model, but when prices are falling the lagged response is much more protracted. As a single measure of the asymmetry, the ratio $(1 - \lambda_2)/(1 - \lambda_1)$ is quite informative, which in this case is .35. This is equal to the corresponding ratio of the

Table 1. Regression Equations for Aggregate Crop Output Response in the United States (1914-51)

Equation	1	2	3
Intercept	–.008 (.018)	–.005 (.008)	–.006 (.007)
Weather Index	.234 (.045)	.223 (.048)	.220 (.047)
Trend	.00730 (.00126)	.00831 (.00566)	.00949 (.00920)
Price Level	.217 (.055)	.241 (.097)	.279 (.136)
Price Increase (β_1)	–	–.196 (.078)	–.207 (.121)
Lag Parameter (λ_1)	–	.707 (.191)	.772 (.160)
Price Decrease (β_2)	–	–.155 (.099)	–.207 (.121)
Lag Parameter (λ_2)	–	.888 (.162)	.921 (.121)
Price Change	–.153 (.055)	–	–
Lag Parameter	.769 (.147)	–	–
$\bar{R}^2$	.862	.858	.861
Standard Error Estimate	.0316	.0321	.0318
Durbin Watson	2.01	2.06	2.02

Note: The numbers in parentheses are asymptotic standard errors and estimates of the initial condition parameters are not reported.

coefficients on p_{t-2} in this application where $\hat{\beta}_1 = \hat{\beta}_2$. The same ratio for higher order lags is increasing proportional to powers of λ_2/λ_1, which can be seen by examining (13).

The orthodox conclusion from the statistical results in Table 1 would be that the asymmetric model is not supported by the data because the hypothesis that $\beta_1 = \beta_2$ and $\lambda_1 = \lambda_2$ cannot be rejected at a reasonable level of significance. But on the other hand if we strongly believe in the asymmetric model because of theoretical considerations, we might take the position that the data do not contain sufficient information on the question to provide a definitive answer. An informative set of data would allow us to estimate the more general asymmetric model with sufficient precision that the point estimates of λ_1 and λ_2 would be quite close to one another with respect to practical interpretation if the hypothesis $\lambda_1 = \lambda_2$ cannot be rejected. The results in Table 1 show estimates of λ_1 and λ_2 which imply much asymmetry, but the precision is very weak, and we would have to conclude that the general model encompassing asymmetry is an over-parameterization with respect to this data set. Nevertheless, we need to remind ourselves of the tenuous nature of any conclusions about the existence of asymmetries in supply response.

At the close of the section on the Wolffram method, the problem of asymmetry being confounded with trended variables was noted. Although the problem is not obvious in the results of Table 1 except for relatively low precision in the trend and price variables, another specification where the years for two large residuals were dummied out of the estimation (1934 and 1936) gave results in which the confounding was extreme. Apparently these two observations provide a disproportionate amount of information on asymmetries in response, if they exist.

One possible explanation for symmetry in response is that the stickiness in adjustments to increases in returns has been neglected by those enamored with asset fixity on the downside. Under the simultaneous conditions of output price uncertainty, limitations of capital goods suppliers to respond to sudden increases in demand, and possible labor shortages during a growth phase, the lags in response might be quite balanced between rising and falling returns. Uncertainty of future prices is especially important when long term financial commitments must be made to finance increased output.

Dynamic Modeling of Supply Response

A Conceptual Model

Some fairly recent research on duality for a dynamic theory of the firm illustrates how measures of capital stocks and other quasi-fixed factors of production enter into product supply functions (Epstein, Taylor). Such a supply function jointly with a system of difference (or differential) equations describing the dynamic behavior of quasi-fixed factors constitutes a logically consistent model of dynamic supply response. Rather than pursuing a rather general analysis of the dynamics of supply couched within such a system of equations, a simple model with only one quasi-fixed factor is used to illustrate the primary aspects of apparent irreversibilities in supply response.

Although price expectations are an important consideration in producer's behavior, a simplifying assumption is made that last year's price, or a futures market contract price, is taken as expected price. Capital investments in agriculture tend to follow movements of output prices for the obvious reason that expansion usually requires more capital goods to maintain efficiency, but also because a progressive income tax makes the net cost to a farmer relatively less during a high income year and vice versa. The following linear equation for an aggregate measure of capital stocks is used to illustrate the way in which fixed asset theory enters the supply function,

$$(23) \quad x_t = \lambda x_{t-1} + \gamma p_{t-1} + \alpha,$$

where x and p are capital stocks and expected output price, respectively. (The notation here is unrelated to previous sections.) Unknown parameters are λ, γ, and α, the first being associated with capital depreciation. An exponential decay relation is assumed to approximate annual depreciation at the rate $(1 - \lambda)$ which yields a survival proportion equal to λ. The linear term in p_{t-1} measures the way new investment responds to product price; price of capital is suppressed in the intercept for simplicity (constant price of x).

A simple equation to illustrate supply response is

$$(24) \quad q_t = \beta_0 + \beta_1 p_{t-1} + \beta_2 x_t + \beta_3 (p_{t-1} x_t),$$

where q_t is output and the $\{\beta_i\}$ are unknown parameters. The last term in (24) reflects an interaction between price and capital stocks with respect to marginal effects of either variable on output response. Without this term, capital stocks would affect only the level of output and not the marginal effect of price movements on output. This latter relationship between marginal response to price and capital stock levels is the essence of fixed asset theory as it relates to agricultural supply.

Equations (23) and (24) together comprise a dynamic model of supply. A priori reasoning suggest λ, γ, β_1, and β_2 are positive; λ must be less than one for stability; and fixed asset theory implies β_3 is negative. Taking the partial derivative of q_t in (24) with respect to p_{t-1} makes the latter assertion clear.

$$(25) \quad \begin{aligned} \partial q_t/\partial p_{t-1} &= \beta_1 + (\beta_2 + \beta_3 p_{t-1})(\partial x_t/\partial p_{t-1}) + \beta_3 x_t \\ &= (\beta_1 + \beta_2\gamma) + \beta_3\gamma p_{t-1} + \beta_3 x_t. \end{aligned}$$

When x_t is large and $\beta_3 < 0$, marginal response to price is relatively low and vice versa because of the last term in (25). This structure makes sense only if there are limiting factors to production, such as declining quality of agricultural land as expansion takes place or external diseconomies. It is shown below that the long-run response function associated with (23) and (24) is concave, and the concavity results from $\beta_3 < 0$.

Since partial derivatives of analytic functions are the same in absolute value for either direction of the infinitesimal change, the relationship in (25) which takes x_t as given is "reversible". But observation of a time series on prices and quantities will give the impression of an irreversible supply response to price. A sequence of price increases starting from a relatively low value of x will show large increases in output as capital stocks and output grow together since both of the linear terms in (24) will be increasing i.e., $\beta_1 p_{t-1}$ and $\beta_2 x_t$. As prices peak and begin to fall, net investments in capital stocks will go to zero and capital stocks will start falling with a lag behind the falling price of output because of the dynamic structure in (23). Capital stocks their equilibrium state for any contemporaneous price as prices were rising, and there will be a rather flat time path for stocks as prices peak and turn downward, ultimately followed by declining stocks. In this period of transition, (25) indicates a relatively small marginal response of output to the recently experienced falling prices. But as capital stocks purchased during the boom depreciate to a low level, output will become more responsive to price changes in either direction.

The phenomenon described above can be clarified by solving the difference equation for x_t in (23) and substituting the results into (24). This gives an equation for output response to current and lagged prices, i.e., a distributed lag model. Solution of (23) by sequentially substituting the same equation for $x_{t-1}, x_{t-2}, \ldots$ on the left hand side into the right hand side yields

$$(26) \quad x_t = \alpha/(1-\lambda) + \gamma(p_{t-1} + \lambda p_{t-2} + \lambda^2 p_{t-3} + \ldots).$$

The intercept was simplified by the properties of geometric progressions, $1 + \lambda + \lambda^2 \ldots = 1/(1 - \lambda)$. Substitution of (26) into (24) yields

$$q_t = \beta_0 + \beta_1 p_{t-1} + (\beta_2 + \beta_3 p_{t-1}) \left[\alpha/(1-\lambda) + \gamma(p_{t-1} + \lambda p_{t-2} + \lambda^2 p_{t-3} + \ldots)\right].$$

Regrouping terms gives

$$\begin{aligned}(27) \quad q_t = {} & [\beta_0 + \beta_2 \alpha/(1-\lambda)] + [\beta_1 + \beta_3 \alpha/(1-\lambda)]p_{t-1} \\ & + \beta_2 \gamma(p_{t-1} + \lambda p_{t-2} + \lambda^2 p_{t-3} + \ldots) \\ & + \beta_3 \gamma p_{t-1} (p_{t-1} + \lambda p_{t-2} + \lambda^2 p_{t-3} + \ldots),\end{aligned}$$

where primary interest is in the last group of terms which is a sequence of terms involving cross-products, $p_{t-1}\, p_{t-i}$, $i = 2, 3, \ldots$. These are the results of the interaction term, $\beta_3(p_{t-1}x_t)$ in (24). Recall β_3 and γ are assumed negative and positive respectively so that each of the lagged price, cross-product terms in (27) is negative. It is noted that the linear lag operator cannot be applied to the last term in (27).

Taking the partial derivative of q_t in (27) with respect to p_{t-1} yields

$$(28)\quad \partial q_t/\partial p_{t-1} = \left[\beta_1 + \beta_2\gamma + \beta_3\alpha/(1-\lambda)\right] + \beta_3\gamma p_{t-1} + \beta_3\gamma(p_{t-1} + \lambda p_{t-2} + \lambda^2 p_{t-3} + \ldots),$$

where it is seen that the marginal response is dependent on a weighted sum of past prices, and the weights decline exponentially from one to a limit-value of zero. Therefore, marginal output response will tend to be relatively small when recent historical prices have been relatively large, and vice versa, which will give the appearance of supply irreversibilities for a typical time series pattern on output prices.

If p is held fixed at some value, say p_*, then x in (23) will approach a limit, x_*. Likewise, if p and x in (24) are set equal to p_* and x_*, the implied equilibrium for output, q_*, is obtained. These long-run equations for capital stocks and output are

$$(29)\quad x_* = (\alpha + \gamma p_*)/(1-\lambda)$$

$$(30)\quad q_* = \left[\beta_0 + \beta_2\,\alpha(1-\lambda)\right] + \left[\beta_1 + (\beta_2\gamma + \beta_3\alpha)/(1-\lambda)\right]p_* + \left[\beta_3\gamma/(1-\lambda)\right]p_*^2.$$

Quantity supplied in long-run equilibrium is a concave quadratic function of price.

The main purpose of this simple model has been to illustrate how apparent irreversibilities arise in typical time series data, but the actual phenomenon can be modeled with ordinary "reversible" functional relationships if the state variables describing the entire dynamic process are incorporated. One is then working with a system of dynamic equations, one for each state variable entering the supply function plus the supply equation itself. Provided that the dynamic system is stable, a long-run equilibrium will exist for price held constant into perpetuity.

Although the correct approach is clear conceptually, many problems exist in practical modeling. Capital goods are extremely heterogeneous and too numerous to use without aggregation, but aggregate measures have obvious limitations. Geometric depreciation is often used for measures aggregated across firms and separate capital items, but is at best an approximation. Technological change results in the quality of capital inputs changing over time so that even a disaggregated capital stock variable becomes ambiguous. Then frequently all the above difficulties are moot because data are not available anyhow, or at least not in a form needed and with sufficient accuracy to be useful. Consequently, research workers usually try to model supply response without jointly modeling the behavior of capital stocks.

Empirical Models

The simple model defined by (23) and (24) was reduced to a single distributed lag equation, (27), expressing quantity supplied as a function of lagged output prices, which avoids estimation of a capital stock equation explicitly. The depreciation parameter λ is identified, but other individual parameters from (23) and (24) are not. Nevertheless, a long-run supply equation is identified from estimates of the composite coefficients in (27), i.e., the parameters in (30). A short-run relationship from estimates in (27) would have to be interpreted as a reduced form of sorts where capital stocks are tacitly embedded in the lag structure on output price.

Since (27) is conditionally linear for a given value of λ, least squares estimates could be computed by a sequence of linear regressions and a search over $0 < \lambda < 1$. The infinite series in weighted lagged prices can be truncated at the last sample observation and the presample part replaced with an unknown parameter.[5] Letting ϕ_0 denote the presample series,

$$p_{t-1} + \lambda p_{t-2} + \lambda^2 p_{t-3} + \ldots = p_{t-1} + \lambda p_{t-2} + \ldots \lambda^{t-1} p_1 + \lambda^t \phi_0,$$

where it is seen that the parameter ϕ_0 enters linearly for given λ. Nonlinear least squares could also be used for estimation, but use of analytical derivatives with respect to the parameters is tedious. The better algorithms using numerical approximations for derivatives are another possibility.

The structure in (27) reflects nonadditivity in the lagged response to prices, which is an essential characteristic to deal with fixed asset theory in supply response. A large family of models exhibiting nonadditive lags is provided by nonlinear difference equations in the dependent variable of the regression. In order to avoid confounding the disturbance term with the dynamic response, it is better to specify these as "nonstochastic" difference equations. For example, a simple model similar to (24) written as a statistical equation is

$$(31) \quad q_t = \alpha + \beta p_{t-1} + \gamma E(q_{t-1}) + \delta p_{t-1} E(q_{t-1}) + u_t,$$

where u_t is a disturbance term and $E(q_t) = q_t - u_t$. Taking expectations of both sides of (31) yields a nonstochastic difference equation as the mean for the regression, while a more common model with q_{t-1} on the right hand side would have a mean conditional on q_{t-1}. (See LaFrance and Burt for application of an additive version of (31), i.e., $\delta = 0$, to aggregate U.S. farm supply.)

One could interpret (31) as an approximation to (24) where x_t is replaced by $E(q_{t-1})$. Since prices tend to show a persistence in their level over time (positively autocorrelated) this smooth path would tend to make $E(q_{t-1})$ a sensible index for aggregate capital stocks. A natural generalization of (31) is to use higher order polynomial terms in the two variables, p_{t-1} and $E(q_{t-1})$, but one may encounter parameter estimates which give implausible long-run supply functions. The author has obtained some promising preliminary results for U.S. aggregate farm supply using a functional form which is the product of a convex function in p_{t-1} (with an asymptote) and a quadratic in $E(q_{t-1})$. The implied long-run supply equation is a sigmoid curve with an asymptote for maximum production. However, the statistical precision leaves much to be desired, somewhat like the empirical results reported earlier for the asymmetric model of U.S. crop supply.

Concluding Remarks

The importance of either directly, or indirectly, introducing nonadditive distributed lag response into empirical supply functions should not be exaggerated, but the allowance for purely additive dynamic response in a rather general way can hardly be overemphasized. The smoothness of economic time series frequently allows good approximations to dynamic response without generalizing to nonadditive lags. This does not mean that the "true" response is strictly additive, but within the limitations of the implicit experimental design imposed by the smooth time series data, often only an additive approximation is operational. In the context of dynamics, additivity is analogous to linearity in static models where we readily accept linearity, within the family of monotonic transformations, as about the limit in complexity which our economic data can support statistically.

In some applications, like milk supply response where a primary state variable is the dairy herd, it is feasible to model the dynamic structure of the state variable jointly with the commodity supply equation (see LaFrance and De Gorter for a study of the U.S. dairy industry). On the other hand, it is likely to be infeasible to model the dynamic adjustments in capital stocks and quasi-fixed labor in estimating an aggregate supply function for U.S. agriculture. Labor adjustments are too dependent on transitory factors of the economy which are not stationary enough for parameter estimation, and the basic measurement problems in data on capital stocks probably cancel out the disadvantages of going to a reduced form supply function. It would appear that an allowance for nonadditivity is more important as the level of aggregation increases because of less flexibility in the usage of capital items. Crops which tend to be grown in rotations for disease, insect, and weed control, such as corn and soybeans in the corn belt, are less apt to require a nonadditive lag specification than crops requiring specialized equipment in monoculture production, e.g., wheat in some parts of the Great Plains.

The profession's infatuation with static duality theory as a basis for empirical modelling of supply response is most unfortunate because no attempt is made to identify a dynamic structure which approximates short-run adjustments in time series data. Taylor's recent results for stochastic, dynamic duality suggest that a primal approach is likely to be the more fruitful in modeling time series data because one would be less inclined to choose a rigidly parameterized model which encompasses a very narrow family of hypotheses. In the author's opinion, our economic theory of the firm with presumption of super-rationality under optimization, contrasted with the nature of the data on which an econometric model is based, makes it questionable to let the theory dictate the model except in a rather general way. The data are aggregates of heterogeneous resources: capital is a mix of many separate items of different vintages and ages; labor and management reflect all the diversity of the human species; land combined with natural climate is equally heterogeneous and the amount of land of a given quality is fixed.

If the implication of this assessment of the setting in which supply functions are estimated is that the estimation process must be relatively more empirical than many recent studies would suggest, so be it. That is not to say theory has no place in specification of dynamic supply response; it is a matter of relative weight and how presumptuous the analyst should be in the amount of detail economic theory can provide about the structure of aggregate time series data. As applied economists, we were told by the experts for a decade or two that economic data are so weak that theory must be used to determine the structure within a tightly specified model with very few parameters. Then during the last decade we have been told by the time series enthusiasts that many of our devices to achieve the closely structured and identified models border on being ludicrous with respect to the detailed knowledge assumed. The most fruitful approach would appear to be a compromise between these two extreme viewpoints by letting the data largely determine details of the dynamic structure, but using economic theory to impose constraints on the equilibrium behavior of the dynamic system, as well as providing as much other information (but possibly misinformation) as would appear prudent within the context of a particular application.[6] Certainly, economic theory would always be used to provide candidates for explanatory variables in a regression equation.

no 11/14/89 OB-7.0

Footnotes

[1]This type of irreversible regression equation has also found application in the analysis of price margins in the agricultural processing sector (Kinnucan and Forker). An application in macroeconomics which preceded Wolffram was Thurow's study of unemployment dynamics.

[2] Traill, Colemen, and Young argued that the long-run response should be reversible although the short-run response might exhibit irreversibilities. Their discussion was in the context of appropriate distributed lag models, but they did not seem to be concerned about the ambiguity of an equilibrium state for a fixed price level. Their modification of the Wolffram technique has special problems of its own (see LaFrance and Burt).

[3] One would expect a rather complex relationship between income and price response with habit formation. A period of increased consumption caused by greater income would alter behavior with respect to prices, which implies an interaction between lagged income and current price response of the consumer. There is also the influence on consumption of prices of related goods; see Young for a study of these issues and an application to coffee demand.

[4]This form of irreversible model could be used in marketing margin studies by measuring prices and costs in deflated instead of nominal dollars, thus providing the opportunity to impose more structure, viz., a unique steady-state for any given level of price regardless of transient changes in prices.

[5] This innovation was first used by Klein and can be found in many text books under the discussion of geometric lags or more general rational lags (Theil, Kmenta).

[6] The author's position here has been heavily influenced by the applied research philosophy of David Hendry (see Hendry, Hendry and Richard, and Hendry and Wallis).

References

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

Burt, Oscar R. "Estimation of Distributed Lags as Nonstochastic Difference Equations." Staff Paper 80-1, Economics Department, Montana State University, Bozeman, 1980.

Burt, Oscar R. "Testing for the Length of a Distributed Lag with a Differencing Transformation." Economics Letters. 29(1989):221-23.

Epstein, L. "Duality Theory and Functional Forms for Dynamic Factor Demands." Review of Economic Studies. 48(1981): 81-95.

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

Harvey, A.C. The Econometric Analysis of Time Series. New York: John Wiley and Sons, 1981.

Hendry, David F. "Econometrics — Alchemy or Science." Economica. 47(1980): 387-406.

Hendry, David F. and Jean-Francois Richard. "On the Formulation of Empirical Models in Dynamic Econometrics." Journal of Econometrics. 20(1982):3-33.

Hendry, David F. and Kenneth F. Wallis. eds. Econometrics and Quantitative Economics. Oxford: Basil Blackwell, 1984.

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

Johnson, Glenn L. "The State of Agricultural Supply Analysis." Amer. J. Agr. Econ. 42(1960): 435-52.

Kinnucan, Henry W. and Olan D. Forker. "Asymmetry in Farm-Retail Price Transmission for Major Dairy Products." Amer. J. Agr. Econ. 69(1987): 285-92.

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

Kmenta, Jan. Elements of Econometrics. 2nd ed. New York: Macmillan, 1986.

LaFrance, Jeffrey T. and Oscar R. Burt. "A Modified Partial Adjustment Model of Aggregate U.S. Agricultural Supply." West. J. Agr. Econ. 8(1983): 1-12.

LaFrance, Jeffrey T. and Harry de Gorter. "Regulation in a Dynamic Market: The U.S. Dairy Industry." Amer. J. Agr. Econ. 67(1985): 821-32.

Taylor, C. Robert. "Stochastic Dynamic Duality: Theory and Empirical Applicability." Amer. J. Agr. Econ. 66(1984): 351-57.

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

Thurow, Lester. "The Changing Structure of Unemployment: An Econometric Study." Review of Economics and Statistics. 47(1965): 137-49.

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

Tweeten, Luther G. and C. Leroy Quance. "Positivistic Measures of Aggregate Supply Elasticities: Some New Approaches." Amer. J. Agr. Econ. 51(1969): 342-52.

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

Young, Trevor. "Modelling Asymmetric Consumer Responses, with an Example." J. Agr Econ. 31(1980): 175-86.