Level of Structural Aggregation and Predictive Accuracy of Milk Supply Response Estimates

Blair J. Smith and Donald R. Scott

Milk supply response was estimated for Pennsylvania using three different levels of structural aggregation. The base level involved the estimation of milk production in a single equation. Under the second method, production was the product of two equations: milk per cow and number of milk cows. The third method factored production into three equations: milk per cow, number of dairy farms, and number of cows per farm. As expected, the greater the degree of disaggregation the more was learned about the structural aspects of milk production. At the same time, predictive accuracy generally decreased, but the differences among models was slight.

Introduction

Most positivistic methods of estimating milk supply response can be categorized into one of two groups: direct and indirect. In the direct method, total milk production is specified as the dependent variable of a multiple regression equation. This equation is then estimated and the results are interpreted as an estimate of milk supply response. The indirect method involves the estimation of two equations whose dependent variables are the number of dairy cows and milk production per cow, respectively. The product of these equations is a second method of estimating milk supply response. Both of these methods have been used successfully to model milk supply response. For instance: Criner; Elterich and Masud; and Prato have had success with the indirect approach while Fallert and Hallberg; Chen et al.; and Stammer have obtained logical response estimates from the direct method.

A third method of estimating milk supply is presented in this article. It is an expanded-indirect approach, wherein milk production is factored into the three equations: number of commercial dairy farms; number of dairy cows per farm; and milk production per cow. Serial multiplication of these three equations results in a third estimate of milk supply response. If this approach has been tried by other researchers, the present authors are not aware of any such efforts.

The primary objective of the study on which this report is based was to estimate milk supply response in Pennsylvania using the traditional (direct and indirect) methods, along with the expanded-indirect approach. The three methods of estimating milk supply response were compared to determine which one predicted most accurately. A priori, the direct method was expected to predict milk supply response more accurately than the indirect and expanded-indirect approaches. The indirect and expanded-indirect approaches, by their disaggregating procedures, increase the probability of error in the estimation process. At the same time, it was hoped, the more highly the milk supply response equation was factored, the better one could understand the structural relationships or components of the milk production sector of the dairy industry. Detail on methodology, data, and results beyond that reported in this article may be found in Scott.

Methodology and Data

Five variables had to be estimated in order to use all three methods of estimating milk sup-

Blair J. Smith is associate professor of agricultural economics and Donald R. Scott is a former graduate student at The Pennsylvania State University. Journal Paper No. 7214, Pennsylvania Agricultural Experiment Station.

ply response. These five variables are: number of commercial dairy farms (CDF), number of dairy cows per farm (CPF), milk production per cow (MPC), total number of dairy cows (NDC), and total milk production (PRO). These variables were specified as dependent variables in five separate regression equations. The equations were estimated with the use of the third-order autoregressive least squares procedure in the Statistical Analysis System (SAS) software package.

The data used in the analysis are for Pennsylvania for the years 1967–81, and all but two variables, number of commercial dairy farms (CDF) and number of replacement heifers greater than 500 pounds (HEF), were available on a calendar quarter basis. Quarterly values for CDF and HEF were developed from annual data by interpolation. The basic sources for all of the data are the Pennsylvania Department of Agriculture and the U.S. Department of Agriculture. Some variables, such as CPF and MAR (profit per cow), were calculated by transforming data from the above sources. The variables included in the final estimating equations are as follows:

BFP = prices received for beef cows (¢/cwt.)
CDF = number of commercial dairy farms (annual data)
CPF = number of dairy cows per farm
MPC = milk production per cow (pounds)
NDC = number of dairy cows (thousands of head)
PRO = milk production (millions of pounds)
DUM = 0, 1 dummy variables to act as intercept shifters for CDF, CPF, MPC, NDC, PRO. These dummy variables are set up as follows:

Dependent variable	Years Included in Dummy Variable: 1	0
CDF	1967–72	1973–81
CPF	1967–70	1971–81
MPC	1967–73	1974–81
NDC	1967–73	1974–81
PRO	1967–70	1971–81

FDP = price of 16 percent protein mixed dairy feed ($/ton)
HEF = number of replacement heifers greater than 500 pounds (annual data)
MKP = price of milk sold to plants and dealers (¢/cwt.)
CPM = cost of producing milk (¢/cwt.)[1]
MAR = profit per cow ($) = (MKP − CPM) × MPC
MFR = MKP/FDP (milk-feed price ratio)
QT2 = 1 for second calendar quarter, 0 otherwise
QT3 = 1 for third calendar quarter, 0 otherwise
TIM_i = a time-trend variable incremented quarterly from i = 1 for the first quarter 1967 through i = 60 for the fourth quarter 1981
SLO = DUM × TIM_i = interaction terms to act as slope shifters on the time variable (TIM) in the CDF, CPF, MPC, NDC, and PRO equations.
t = indicates the variable is not lagged, t − 1 indicates a one-quarter lag, t − 2 a two-quarter lag, etc.

Results

Plots of the dependent variables under investigation (CDF, CPF, MPC, NDC, and PRO) revealed significant changes either in direction of trends or in rates of change in some of the variables during the period 1970–73. The explanatory variables DUM and SLO were employed to account for these changes, but no attempt was made to ascertain why such changes occurred. Both DUM and SLO were found to have significant impacts on CDF, NDC, and PRO (Table 1). Neither DUM nor SLO were significant in either the MPC or CPF estimating equations.

Table 1 shows the number of commercial dairy farms in Pennsylvania was also a function of the milk-feed price ratio and time. The lag in response to changes in the milk-feed price ratio was less (within 3 quarters) than the 4 to 8 quarters that was initially expected. The reason for the initial expectation is two-fold. First, the amount of investment required to enter into dairying imposes a substantial barrier for most potential entrants, and they

[1] Costs of producing milk in Pennsylvania are computed quarterly from data supplied each month by Pennsylvania dairy farmers. The estimates of cost are meant to include all costs incurred in the production of milk except allowances for the operator's management and equity capital. A full description of the procedures used to estimate cost of production can be found in Smith and Sedlak.

Table 1. Regression Results for Number of Commercial Dairy Farms (CDF), Number of Dairy Cows Per Farm (CPF), Milk Production Per Cow (MPC), Number of Dairy Cows (NDC), and Total Milk Production (PRO), Pennsylvania, 1967–81

$CDF = 17290.16^{*} + 4586.26\ DUM^{*} + 108.73\ MFR_t^{*} + 102.60\ MFR_{t-2}^{**} - 187.82\ SLO^{*} - 87.50\ TIM^{*}$	$\bar{R}^2 = 0.96$
$CPF = 53.2800^{*} - 0.0003\ BFP_t^{***} - 0.0010\ CDF_t^{*} + 0.5000\ HEF_{t-1}^{*} + 0.0180\ MAR_{t-1}^{*} + 0.0680\ TIM^{**}$	$\bar{R}^2 = 0.98$
$MPC = 1513.93^{*} + 43.40\ MFR_{t-2}^{*} + 45.63\ MFR_{t-6}^{*} + 205.89\ QT2^{*} + 58.66\ QT3^{*} + 13.64\ TIM^{*}$	$\bar{R}^2 = 0.91$
$NDC = 647.52^{*} - 0.005\ BFP_t^{*} + 62.980\ DUM^{*} - 0.098\ FDP_t^{***} + 0.040\ MKP_t^{*} + 0.030\ MKP_{t-3}^{*} + 0.040\ MKP_{t-5}^{*} - 2.720\ SLO^{*}$	$\bar{R}^2 = 0.95$
$PRO = 244.01^{*} - 416.82\ DUM^{*} + 29.52\ MFR_{t-6}^{*} + 1.76\ NDC_t^{*} + 44.29\ QT2^{*} + 15.28\ QT3^{*} + 14.45\ SLO^{*}$	$\bar{R}^2 = 0.93$

*, **, and *** indicate the corresponding coefficients are statistically significant at the 1, 5, and 10 percent levels, respectively.

would like to feel reasonably assured that milk prices will remain favorable during the investment pay-back period before committing such large sums to the production of milk. Second, the fixed-asset nature of dairy farming tends to prolong the exodus of dairy farms from the industry, even when the relationship between milk prices and costs of production becomes unfavorable.

The number of dairy cows per farm in Pennsylvania was discovered to be a function of beef prices, number of commercial dairy farms, number of heifers, profit per cow, and time. It is evident from Table 1 that the number of dairy cows per farm responds very quickly (within 2 quarters) to changes in the aforementioned variables. This result was not expected since it takes about 2–3 years to raise replacement cows, although by adjusting the rate of culling more rapid changes in cows per farm are possible.

Milk production per cow in Pennsylvania was found to be a function of the milk-feed price ratio, seasonality, and time. The lengths of lags on the milk-feed price ratios are difficult to explain. Changes in milk production per cow are not likely to occur until the dairy farmer is able to implement a revised feeding program and/or substitute higher-producing cows for lower-producing ones. If the quality of feed is fixed by virtue of the fact that much of it is home-grown, and if the herd is currently being fed to maximize production from that feed, the lags in response may be rational. The coefficient on the seasonality variable, QT2, verifies the fact that milk production per cow normally peaks in the spring months.

The number of dairy cows in Pennsylvania is shown to be a function of beef prices, feed prices, and milk prices. The length of lag on the price of milk (up to 6 quarters) was a bit less than expected since it takes roughly 10 quarters to raise additional dairy cows. On further reflection, however, it was realized that changes in culling rates, if circumstances warranted, would alter the length of the lag. Beef and feed prices showed their expected inverse relationship with the number of dairy cows.

Total milk production in Pennsylvania was estimated to be a function of the milk-feed price ratio, number of dairy cows, and seasonality. Milk production per cow influences total milk production. Therefore, factors affecting milk production per cow should also influence total milk production. The appearance of the milk-feed price ratio, lagged six quarters in both the milk production per cow and total milk production equations, supports this statement. The positive influence of the number of dairy cows on total milk production is fairly obvious. The coefficient on the QT2 variable confirms that total milk production peaks in the spring along with production per cow.

Predictive Accuracy of the Three Models

The three methods of estimating milk supply response in Pennsylvania were compared to determine which one predicts with the most accuracy. Table 2 compares the 1982 predicted levels of milk production with the actual levels of production. All three methods predicted production with less than three percent error. Surprisingly, milk production was predicted most accurately by the indirect method. This does not lend support to the previously held notion that milk supply response

Table 2. Predicted and Observed Values of Milk Production in Pennsylvania, 1982

Method of estimation	Milk production		Difference	
	Predicted level	Observed level	Pounds (millions)	Percent of observed value
	million pounds			
Direct[a]	9,437	9,264	+ 173	+ 1.87
Indirect[b]	9,100	9,264	− 164	− 1.77
Expanded-indirect[c]	9,046	9,264	− 218	− 2.35

[a] Direct method: PRO = PRO
[b] Indirect method: NDC × MPC = PRO
[c] Expanded-indirect method: CDF × CPF × MPC = PRO

would be predicted most accurately using the direct method. However, the difference between the prediction errors for the direct and indirect approaches was small. As expected, however, the direct and indirect methods predicted milk supply response more accurately than the expanded-indirect method. Again, this is attributed to the error introduced into the estimation procedure when milk supply response is factored into its component parts.

Summary and Conclusions

The primary objective of this study was to estimate and explain milk supply response in Pennsylvania using the two traditional positivistic methods, direct and indirect, plus an expanded-indirect approach. It was hypothesized that the expanded-indirect approach would provide more information about milk supply response in Pennsylvania than either the direct or indirect method. The results from this study support this hypothesis. Under the direct method, total milk production was found to be a function of the milk-feed price ratio, the number of dairy cows, and seasonality.

The indirect method disaggregated milk supply response into the number of dairy cows and milk production per cow. The number of dairy cows was found to be a function of the prices of beef, feed, and milk. Milk production per cow was discovered to be a function of the milk-feed price ratio, seasonality, and time. Thus, the indirect method provides somewhat more structural information on milk supply response in Pennsylvania than the direct method.

The expanded-indirect approach takes the indirect approach one step further by disaggregating the number of dairy cows into the number of commercial dairy farms and the number of dairy cows per farm. The number of commercial dairy farms was found to be a function of the milk-feed price ratio and time. The number of dairy cows per farm was found to be a function of beef prices, the number of commercial dairy farms, the number of heifers, profit per cow, and time. This approach introduces two new variables, number of heifers and profit per cow, that do not appear in the other two models. Hence, it can be concluded that the expanded-indirect approach does provide additional structural information not already produced by the direct and indirect approaches. The question of whether to use this new approach depends on the goals or purposes of the study.

It was also hypothesized that the direct method would predict more accurately than the indirect and expanded-indirect methods. Table 2 revealed that the indirect approach predicted most accurately. The prediction errors for all three methods, however, were rather small. Thus, it cannot be concluded that any one of the three methods has greatly superior predictive capability.

References

Chen, D., R. Courtney, and A. Schmitz. "Polynomial Lag Formulation of Milk Production Response." *American Journal of Agricultural Economics* 54(1972):77–83.

Criner, G. K. "An Econometric Model of Milk Production in Maine." The University of Maine at Orono, 1983 (Mimeographed).

Elterich, G. J. and S. M. Masud. *A Distributed Lag Model of Milk Supply Response for Delaware.* The University of Delaware Agr. Exp. Sta. Bull. No. 425, Nov. 1978.

Fallert, R. F. and M. C. Hallberg. *Policy Simulation*

Model for the United States Dairy Industry. The Pennsylvania State University Agr. Exp. Sta. Bull. No. 805, Jan. 1976.

Pennsylvania Department of Agriculture. *Crop and Livestock Annual Summary, 1981*. The Pennsylvania Crop Reporting Service (CRS), and the Statistical Reporting Service (SRS).

———. *Historic Dairy Statistics, 1960–1979*. CRS, SRS, 1980.

———. Special Dairy Report. CRS, SRS (various issues).

Prato, A. A. "Milk Demand, Supply, and Price Relationships, 1950–1968." *American Journal of Agricultural Economics* 55(1973):217–222.

SAS Institute Inc. *SAS/ETS USER'S GUIDE*: *Econometric and Time-Series Library*, Box 8000, Cary, NC, 27511, 1982 Edition, pp. 187–202.

Scott, D. R. "Estimation of Milk Supply Response in Pennsylvania Using an Autoregressive Least Squares Procedure." Unpublished M.S. thesis, Department of Agricultural Economics and Rural Sociology. The Pennsylvania State University, University Park, PA, May 1984.

Smith, B. J. and W. J. Sedlak. *Procedures Used by the Pennsylvania Crop Reporting Service to Estimate the Cost of Producing Milk*. The Pennsylvania State University, Department of Agricultural Economics and Rural Sociology, A.E. & R.S. 161, September 1982.

Stammer, R. W. "The Estimation of Milk Supply Response from Micro-Level Federal Order Data." In M. C. Hallberg and B. J. Smith (eds.), *Proceedings of a Workshop on the Effects of Alternative Pricing Structure and Pooling Methods on the Dairy Industry in the Northeast, June 1975*. The Pennsylvania State University, Department of Agricultural Economics and Rural Sociology, A.E. & R.S. 120, Oct. 1975.

U.S. Department of Agriculture. *Agricultural Prices*. Crop Reporting Board (CRB), SRS, (various issues).

———. *Cattle and Calves*. CRB, SRS, (various years).