Staff Papers Series

Staff Paper P88-22

July 1988

COMPONENT PRICING OF PRODUCER MILK:
A YIELD-BASED MODEL FOR THE CHEESE INDUSTRY

by

Jay S. Coggins and Jerome W. Hammond

Department of Agricultural and Applied Economics

University of Minnesota
Institute of Agriculture, Forestry and Home Economics
St. Paul, Minnesota 55108

Staff Paper P88-22

July 1988

COMPONENT PRICING OF PRODUCER MILK: A YIELD-BASED MODEL FOR THE CHEESE INDUSTRY*

by

Jay S. Coggins and Jerome W. Hammond**

*Selected paper presented at the Annual Meetings of the American Agricultural Economics Association, July 31 - August 3, 1988, Knoxville, TN.

**The authors are, respectively, a Doctoral Candidate and Professor in the Department of Agricultural and Applied Economics, University of Minnesota.

COMPONENT PRICING OF PRODUCER MILK:
A YIELD-BASED MODEL FOR THE CHEESE INDUSTRY[1]

I. Introduction

In recent years, several studies have addressed the question of how milk producers should be rewarded for the levels of solid components (butterfat (BF), protein (PR), and solids-not-fat (SNF)) in their milk. Most authors seem to agree that producer prices should reflect the value of the final product and the relationship between milk components and product yield (Graf; Hillers, *et.al.*; Ladd and Dunn; Perrin). Discovering this relationship has proven difficult, however.

In this paper, commercial plant data are used to estimate a cheese yield production function. This yield relationship is then used as a basis for constructing a pricing scheme for producer milk which accurately reflects the economic value of producer milk components to a cheese plant. The yield formula is of independent interest to the cheese industry for its ability to accurately predict yield in an individual plant, and its derivation is spelled out in some detail below.

II. Relation to the Literature

Historically, prices paid for producer milk have been established with little regard for variations in its solids-not-fat or protein content, largely because it was difficult to test for protein or SNF until fairly recently. Now, these fractions may be determined quite easily and inexpensively. Traditional milk pricing plans (with a base price which is

---

[1]The authors would like to thank Vic Adamowicz, Rob King, Gerald Nolte, Claudia Parliament, and Willis Peterson for providing helpful comments on an earlier draft of the paper.

adjusted for the milk's butterfat level) have also been criticized because they adhere to an invalid assumption as to the degree of correlation between the butterfat and non-fat content of milk (Froker and Hardin). The relationship between these variables is actually weaker than was often assumed in decades past (Hillers, *et.al.*).

Several writers have proposed pricing schedules which overcome this difficulty by attempting to value milk components based upon final product yield (Ladd and Dunn; Hillers, *et.al.*; Ernstrom). These studies incorporate a cheese yield formula ($CH = (BF(.93)+(C-.1))\cdot(1.09)/(1-W)$; where C is milk casein, CH is cheese yield, and W is cheese moisture (Van Slyke and Price)) to predict cheese yield, a formula which has two drawbacks for the purpose. First, it is not plant specific, but, rather, it assumes a fixed production technology across manufacturers. Second, it is linear in the butterfat and protein variables, so that their substitution elasticity is assumed to be infinite. We doubt that such an assumption is reasonable.

III. The Cheese Yield Function

As economists, we are accustomed, by the nature of our discipline, to working with data generated in a non-experimental fashion. What's more, we are not surprised when these data exhibit important departures from the ideal. Indeed, econometrics developed as a distinct branch of statistics for the very reason that economic data so often display one or another of serial correlation, heteroscedasticity, multicollinearity, etc.

While one may hope, by designing laboratory cheese-making experiments, to collect cheese yield data devoid of them, observations which are recorded on the usual operations in a commercial plant will have these difficulties. Data of this sort are almost certain to be correlated over time, and the variables which may be expected to best predict yield of cheese per unit of milk, namely the milk's butterfat and protein content, are inherently collinear.

Thus, the task of obtaining accurate and reasonable cheese yield formulas at the plant level, using numbers generated by a commercial process, falls naturally but somewhat surprisingly within the usual province of the econometrician. In this section econometric tools are applied to the problem which technical cheese researchers have historically addressed by a mass balance approach (see, e.g. Van Slyke and Price; Lelievre, *et.al.*). There, the milk is carefully measured for all component levels before being made into cheese in a controlled experimental environment. The composition of the cheese is then analyzed, and the resulting yield formula specifies the share of each of the milk components which should be retained in the cheese, and the share which will be included in the whey.

The Data and Functional Form

The data for the analysis were taken from the daily records of a commercial Midwest cheddar plant. Observations on milk volume received, its butterfat, protein, and SNF content, the total volume of cheese produced and its moisture content were obtained for each of 660 days spanning 23 months.

The Cobb-Douglas (C-D) functional form was selected for the yield production function. Finding the SNF variable to be insignificant, for each functional form, in explaining cheese yield, we chose to use the milk fat and protein levels (percentages) as explanatory variables. The dependent variable was the moisture-corrected cheese yield (CH) per hundredweight (cwt) of milk. Upon finding none of the parameter estimates different from zero at the 0.10 level of significance for the translog or quadratic functional forms, attempts to fit the data to these forms were abandoned.

Dealing with serial correlation

As these were daily observations, it was not surprising that a preliminary analysis revealed a distinct seasonal pattern in the milk component levels and the cheese yield values. It was hypothesized that a seasonal dummy variable specification would be an appropriate remedy for this long and slow movement around the predicted relationship. Seasons were specified as three-month intervals, with the spring season beginning on the first day of March. The model of choice, then, in logged form, is

$$\ln (CH) = \beta_0 + \beta_1 \ln (BF) + \beta_2 \ln (PR) + \beta_3 D_S + \beta_4 D_F + \beta_5 D_W + u \qquad (1)$$

where

CH = cheese yield per cwt of milk, corrected to a standard moisture level of 38%,

BF = percentage level of butterfat in the milk,

PR = percentage level of protein in the milk,

u is a disturbance term which is assumed to follow the normal distribution with mean zero and variance matrix $\sigma^2 I_{660}$, and

D_j takes on the value one for observations in season j and zero otherwise; j takes values summer (S), fall (F), and winter (W).

The hypothesis of seasonality was tested by fitting (1) and the Cobb-Douglas form

$$\ln (CH) = \beta_0' + \beta_1' \ln (BF) + \beta_2' \ln (PR) + u \qquad (2)$$

Model (1) allows the intercept parameter to vary across seasons, while the slopes remain fixed; (2) restricts the model to a common intercept term. The F-test for comparing models (1) and (2) indicated at the 0.01 level that the intercept dummy variables should be included in the model. There was some evidence that the slope parameters also move seasonally. That specification led to negative marginal products in some periods, however, and was abandoned. The form of the production function, then, is given by (1). Tables 1a and 1b summarize the results of estimating (1) and (2).

It may be concluded that both models fit the data quite well. In each case, the value of the coefficient on ln(BF) is slightly more than twice the coefficient on ln(PR), a fact which will be important in the pricing formula discussion. The negative coefficient on D_s shows that the yield of cheese is systematically lower during the summer months than in other periods, as would be expected from climactic and milk composition considerations.

Dealing with multicollinearity

It is widely known that as the percentage of butterfat in milk increases or falls, its protein content moves in the same direction (Froker and Hardin). Clearly, this fact, together with our stated intention to use butterfat and protein content as explanatory variables in the production function estimation, gives a hint that collinearity may be a concern. By the singular value decomposition of the 660-by-6 design matrix, the condition number was found to equal 272.37. This far exceeds the rule of thumb cut-off level for assessing damaging multicollinearity of 100 (Belsley, *et.al*).

While multicollinearity does not bias the parameter estimates, it inflates their standard errors and thereby reduces the reliability of the estimates. The oldest technique for remedying multicollinearity is to add observations. With 660 already in the sample, it seems unlikely that this would help very much. In fact, the sheer size of the data set adds to our confidence in the estimates.

The collinearity problem was remedied by implementation of Theil's mixed

Table 1a

RESULTS FROM ESTIMATION OF THE MODEL WITH
SEASONAL DUMMY VARIABLES (Equation (1))

Dependent variable: ln(CH)

Observations:	660	Degrees of freedom:	654
R-squared :	0.790	Rbar-squared :	0.788
Residual SS :	0.171	Std error of est :	0.016
Total SS :	0.815	F(6 ,654)=491.4988	P-value=0.00

Var	Coef.	Std. Coef.	Std. Error	t-Stat	P-Value
Const	1.498093	0.000000	0.054562	27.456770	0.000000
ln(BF)	0.451142	0.486819	0.043458	10.381145	0.000000
ln(PR)	0.203660	0.123135	0.066542	3.060621	0.002209
SUMM	-0.017997	-0.229255	0.002146	-8.387585	0.000000
FALL	0.008790	0.105067	0.002169	4.053292	0.000051
WINT	0.011251	0.131560	0.001977	5.689927	0.000000

Table 1b

RESULTS FROM ESTIMATION OF THE TWO-VARIABLE
COBB-DOUGLAS MODEL (Equation (2))

Dependent variable: ln(CH)

Observations:	660	Degrees of freedom:	657
R-squared :	0.754	Rbar-squared :	0.753
Residual SS :	0.201	Std error of est :	0.017
Total SS :	0.815	F(3 ,657)=1005.0717	P-value=0.00

Var	Coef.	Std. Coef.	Std. Error	t-Stat	P-Value
Const	1.116606	0.000000	0.044081	25.331033	0.000000
ln(BF)	0.642577	0.693393	0.038587	16.652652	0.000000
ln(PR)	0.317856	0.192180	0.068868	4.615429	0.000000

estimation procedure (Theil and Goldberger; Theil, pp. 347-352). Stated briefly, by this technique a researcher introduces non-sample information, consisting of stochastic restrictions on the parameter estimates, into the estimation procedure.

Pure Bayesian techniques may also be used to remedy collinearity, but they are analytically difficult and they demand much of the researcher's subjective notions of the correct model. Ridge regression, another alternative remedy, places very firm and inflexible prior restrictions on the model, and in a way which is difficult to justify from a Bayesian viewpoint. Mixed estimation falls between these two extremes in computational difficulty and flexibility. In fact, the ridge estimator is easily shown to be a special case of the mixed estimator. The reader is referred to the Appendix for a discussion of the technical aspects of mixed estimation and ridge regression.

The prior restriction placed on the model simply requires that the *sum* of the parameter estimates $\hat{\beta}_1 + \hat{\beta}_2$ be "near" 0.65, the value of this sum from the OLS estimation.[2] Thus, in the notation of the appendix, we chose the scalar c = 0.65 and specified R as the 1×6 matrix $R = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}$. It was assumed that the sum $\hat{\beta}_1 + \hat{\beta}_2$ is normally distributed about 0.65, and three levels of "nearness" for this value were specified. These were given by intervals around 0.65 within which we wish to believe with 95% confidence the true value would fall. Under the normality assumption, each of the 95% confidence intervals implies a variance for the sum.

In Cases I - III, respectively, the 95% intervals were (0.55,0.75),

---

[2]Strictly speaking, this prior restriction is not in the spirit of mixed estimation, in which only non-sample information should be included in the restrictions.

(0.60,0.70), and (0.64,0.66). The implied variances are given by 0.002603, 0.00065, and 0.000026 respectively (see Belsley, *et.al.*, pp. 198ff). In Table 2 the mixed estimation results for these three degrees of tightness in prior information are presented.

An informal comparison of the results in that table with Table 1a readily supports the observation that the OLS estimate is quite reliable. More formally, it is encouraging that while the condition number in Case III was reduced by more than 25% from the OLS case, the parameter estimates themselves moved only in the third decimal place. This also suggests that the estimates are indeed quite reliable. The parameter estimates of Case II, reported in Table 2b, were used for purposes of devising the pricing plan.

To conclude this section, two observations on the estimation procedure are noted. First, the data set was of remarkably good quality. Even at the plant level, it is unlikely that a more carefully gathered set of numbers could be obtained. And yet the correlation and collinearity problems were substantial. These problems are inherent, and this work represents the most reasonable method of making use of an imperfect dataset.

Second, the prior information used in the mixed estimation procedure was ad hoc in some ways. However, again, if other alternatives may be advocated for various reasons, this choice of restrictions, as well as any other, offers a sensible and reasonable result.

Let us now turn our attention to the second stated goal – that of devising a producer milk pricing schedule from the production function estimation results of Table 2b.

Table 2a: Case I

Observations:	661	Degrees of freedom:	655
R-squared :	0.967	Rbar-squared :	0.967
Residual SS :	0.171	Std error of est :	0.016
Total SS :	5.268	F(6 ,655)=2.2599E+006	P-value=0.0000
Condition Number:	231.357		

Var	Coef	Std. Error	t-Stat	P-Value
Const	1.500638	0.040495	37.056915	0.000000
ln(BF)	0.451342	0.043330	10.416280	0.000000
ln(PR)	0.201293	0.057174	3.520728	0.000430
SUMM	-0.018048	0.002013	-8.967996	0.000000
FALL	0.008815	0.002137	4.125642	0.000037
WINT	0.011279	0.001935	5.829344	0.000000

Table 2b: Case II

Observations:	661	Degrees of freedom:	655
R-squared :	0.961	Rbar-squared :	0.961
Residual SS :	0.171	Std error of est :	0.016
Total SS :	4.441	F(6 ,655)=2.2599E+006	P-value=0.0000
Condition Number:	209.900		

Var	Coef	Std. Error	t-Stat	P-Value
Const	1.502419	0.026594	56.495496	0.000000
ln(BF)	0.451481	0.043264	10.435418	0.000000
ln(PR)	0.199637	0.049623	4.023114	0.000057
SUMM	-0.018085	0.001915	-9.442860	0.000000
FALL	0.008833	0.002115	4.175800	0.000030
WINT	0.011298	0.001906	5.928966	0.000000

Table 2c: Case III

Observations:	661	Degrees of freedom:	655
R-squared :	0.806	Rbar-squared :	0.804
Residual SS :	0.171	Std error of est :	0.016
Total SS :	0.881	F(6 ,655)=2.2625E+006	P-value=0.0000
Condition Number	201.059		

Var	Coef	Std. Error	t-Stat	P-Value
Const	1.503665	0.007402	203.137422	0.000000
ln(BF)	0.451579	0.043218	10.448855	0.000000
ln(PR)	0.198479	0.043569	4.555492	0.000005
SUMM	-0.018110	0.001844	-9.821051	0.000000
FALL	0.008845	0.002100	4.211756	0.000025
WINT	0.011312	0.001885	6.001346	0.000000

IV. The Milk Pricing Mechanism

As noted earlier, there are some drawbacks to using the traditional plan for determining producer prices for milk used in the manufacture of cheese. In this section, an alternative pricing mechanism is devised which is applicable only to milk used in the manufacture of cheese. It more accurately rewards producers for supplying highly productive milk than does any currently defined plan.

This pricing plan is essentially a weighting scheme for attributing the values of cheese and whey products to the various milk solid components. At its center are the marginal physical products (MPPs) of butterfat and protein in the cheese yield production function.[3] Similar yield functions for the various whey products were not estimated because the relevant data were unavailable. Reported average whey product yields were used to complete the weighting scheme.

The pricing mechanism calculates the value of all four of the final products (cheese (CH), whey cream (WC), whey lactose (WLAC), and whey protein concentrate (WPRO)) which this plant produces. Then, using the production function for cheese and the reported monthly whey product yield figures, the end product values are allocated to each of the milk components: BF, PR, and non-fat, non-protein solids (MSOL). Exogenous series are utilized for the product prices which play a role in the pricing schedule.

Our objective is not to establish or derive the optimal base producer

---

[3]The marginal physical product of butterfat, for example, is the partial derivative of the estimated (exponentiated) production function with respect to BF. Where these expressions are used below, they are evaluated at the plant means for the respective variables.

price level. Doing so would require a formidable accounting exercise which would both complicate the paper and obscure the primary results. Instead, we estimate the total monthly producer remuneration and proceed to derive formulas which determine producer prices and accurately reflect the end product value of producer milk components. The important issue is not the size of the pie, but how it should be divided.

The fundamental objective that leads to the derivation presented here must be stated carefully at the outset. It is to calculate, for each of the solid components (BF, PR, and MSOL) of producer milk, implicit (per pound) prices which reflect the true value of the individual constituents in all end products. We propose that manufacturers pay this constant dollar amount *per unit of BF, PR, and MSOL delivered by individual producers.*

To justify this scheme, it is necessary to assume explicitly that no one producer is large enough relative to the daily plant milk supply to exert a measurable effect on the plant-wide milk component tests. Thus, whatever the butterfat or protein *percentages* in an individual farmer's milk, the value to the plant of *a pound of milk component* (BF, PR, or MSOL) is constant across producers. No value is placed on the water in milk (St-Pierre and Scobie). We also abstract from other milk quality considerations, such as the somatic cell count, etc.

The Pricing Schedule

Let P_m denote the monthly average producer price received for a cwt of milk of 3.5% butterfat content as reported by selected Midwest plants, and let DIFF denote their reported monthly BF differential (the amount by which the pay price varies for each 0.1% of butterfat above or below the base level). If $\overline{BF}$ represents the monthly plant average butterfat level, then

$PP_m = P_m + (\overline{BF} - 3.5) \cdot DIFF$, where PP_m denotes the average producer pay price for milk delivered to this plant in each month of the study period. The total value of funds (PREV) allocated to producer remuneration is calculated by multiplying PP_m by the total milk volume (MILK) in cwts:

$$PREV = PP_m \cdot (MILK). \tag{3}$$

It is this quantity, the total monthly producer revenue, which is divided among all producers according to the set of equations below.

It should be recognized that there is an approximation error in deriving average producer price from the data in this manner. However, it is not central to what follows, and in an application the plant would know the appropriate base price which it could afford to pay (and the correct value for PREV).[4]

Since plant operating and fixed cost data were not available, the processing costs could not be assigned to each of the individual end products. Thus, it was assumed that the processing, handling, and labor costs are proportional to the end product value.

The composition of the whey solids is treated in the following manner. Half of the whey solids yield is known from plant data to be processed into a whey protein concentrate product, and the other half into a lactose concentrate. On a dry matter basis, these contain 24, 57, 17, and 2; and 4, 95, 0, and 1 percent of protein, lactose, ash, and butterfat, respectively.

Given the monthly series for the averages of the variables BF, PR, WC, MSOL, WPRO, WLAC, P_{CH} (the cheese price), P_{WC} (the whey cream price), P_{WLAC} (the dried whey lactose concentrate price), and P_{WPRO} (the dried whey

---

[4]A milk buyer must choose the total amount of revenue available to pay producers in a pay period under any plan. That amount, in an implementation of this schedule, would coincide with PREV.

protein concentrate price), and the composition of the dried whey solids products, the value of the end products resulting from processing one cwt of milk may be calculated. Let Y_i represent the yield of the ith end product produced by one average cwt of producer milk, where in what follows i takes on the values CH, WC, WPRO, AND WLAC. Let V_i be given by Y_i times the product price P_i.[5] Let $VAL = \Sigma_i V_i$ denote the total end product value of one cwt of milk.

To extract the share of the total end product value due to each product i, weights W_i are defined as $W_i = V_i / VAL$, where $\Sigma_i W_i = 1$. Using these weights, the share PS_i of the producer price PP_m which is claimed by each of these products is defined as $PS_i = W_i \cdot PP_m$. Note that by construction $\Sigma_i PS_i = PP_m$.

Finally, the *per pound implicit price* (IP) of the components of producer milk are given by the following set of equations:

$$IP_{BF} = \frac{1}{BF}\left\{PS_{CH}\left[\frac{MPP_{BF}}{MPP_{BF} + MPP_{PR}}\right] + PS_{WC} + (.02)PS_{WPRO} + (.01)PS_{WLAC}\right\} \qquad (4a)$$

$$IP_{PR} = \frac{1}{PR}\left\{PS_{CH}\left[\frac{MPP_{PR}}{MPP_{BF} + MPP_{PR}}\right] + (.24)PS_{WPRO} + (.04)PS_{WLAC}\right\} \qquad (4b)$$

$$IP_{MSOL} = \frac{1}{MSOL}\left\{ (.74)\ PS_{WPRO} + (.95)\ PS_{WLAC} \right\}. \qquad (4c)$$

Note that PS_{WPRO} and PS_{WLAC} are divided among the implicit prices for the protein, butterfat, and solids fractions according to the composition of the whey solids products given above. Eqns. (4) comprise a weighting scheme for attributing product values to milk component levels. The terms involving

[5]It is assumed that whey cream is 35 percent butterfat, and that one pound of whey butterfat yields 1.23 pounds of Grade B butter.

MPPs have a compelling intuitive interpretation. If the BF and PR levels in milk were increased by a tiny increment, then cheese yield would rise by the sum of the marginal physical products. The terms involving MPPs in (4a) and (4b) capture the proportion of this increase which is due to the increase in BF and PR, respectively.

Equations (4) can be used to demonstrate that under this pricing plan, total outlays to producers will equal the funds available for those outlays, PREV. For this result, let BFTOT, PRTOT, and MSOLTOT denote the total butterfat, protein, and MSOL pounds which are received by the plant in a month, let R_{TOT} denote the total revenue paid to producers according to equations (4), and let PREV be as defined above. Let MILK denote the total volume of producer milk used for cheese production in the month, in cwts. Also, denote by use of an overbar the sample mean of a milk component. A little algebra allows us to see that

$$\begin{aligned} R_{TOT} &= IP_{BF} \cdot (BFTOT) + IP_{PR} \cdot (PRTOT) + IP_{MSOL} \cdot (MSOLTOT) \\ &= \left[IP_{BF} \cdot (\overline{BF}) + IP_{PR} \cdot (\overline{PR}) + IP_{MSOL} \cdot (\overline{MSOL}) \right] \cdot (MILK) \\ &= \left[PS_{CH} + PS_{WC} + PS_{WSOL} \right] \cdot (MILK) - PP_{m} \cdot (MILK) = PREV. \end{aligned}$$

Thus, by following the pricing schedule, a cheese manufacturer will pay producers a total value which exactly exhausts the revenue available for that purpose, as calculated independently from marginal analysis or from the usual price formulation process.

An Empirical Application

If producer k's milk contains BF_k, PR_k, and $MSOL_k$ percent of butterfat, protein, and solids not-fat not-protein, respectively, then we calculate the realized producer per cwt milk price (P_k) from eqns. (4) as

$$P_k = IP_{BF} \cdot BF_k + IP_{PR} \cdot PR_k + IP_{MSOL} \cdot MSOL_k. \tag{5}$$

Table 3 presents, for twenty milk PR, BF, and SNF combinations, per cwt producer prices under the proposed pricing plan and the traditional plan. Average 1987 milk and cheese quantities and prices are used to demonstrate the impact of the plan.

Our pricing schedule implies that butterfat is underpriced relative to protein by the traditional plan (see footnote 3 to the table), supporting the argument of St-Pierre and Scobie. In general, producers of high-test milk gain relative to lower-test producers by our plan.

V. Conclusions

In this study, a production function for cheddar cheese manufacture was estimated using factory-level data. This yield relationship was used to develop a pricing schedule that is more equitable than current plans. It is practical for cheese manufacturers, and over time it would, by rewarding the most valuable milk supplies, provide producers with an incentive to adjust production in favor of the more valuable milk component by breeding and feeding their cows for this purpose.

Future research could build upon this idea by deriving similar results for the various dairy product sectors. If end product values were estimated in this way for fluid milk, butter, and non-fat dry milk, these could be merged into a pricing schedule that would be meaningful for the entire dairy industry. The results here could also be used in studies of the profitability of handlers' rerouting milk supplies to various plants in a supply region if in fact the technologies in two or more centrally owned cheese plants are relatively more efficient in utilizing various of the milk components.

Table 3

PRODUCER CWT MILK PRICES FOR TWO PRICING PLANS AND 20 MILK COMPONENT COMBINATIONS[1]

MILK % BFAT	MILK % PROT	MILK % SNF	TRADITIONAL PRICING PLAN[2]	PROPOSED PRICING PLAN[3]
3.20	2.90	8.30	11.2759	10.9758
3.20	2.95	8.35	11.2759	11.0399
3.20	3.00	8.40	11.2759	11.1041
3.20	3.05	8.45	11.2759	11.1682
3.40	3.00	8.40	11.5896	11.5082
3.40	3.05	8.45	11.5896	11.5723
3.40	3.10	8.50	11.5896	11.6365
3.40	3.15	8.55	11.5896	11.7006
3.55	3.05	8.45	11.8249	11.8755
3.55	3.15	8.55	11.8249	12.0037
3.55	3.20	8.60	11.8249	12.0679
3.55	3.25	8.65	11.8249	12.1320
3.70	3.20	8.60	12.0602	12.3710
3.70	3.25	8.65	12.0602	12.4351
3.70	3.30	8.70	12.0602	12.4993
3.70	3.35	8.75	12.0602	12.5634
4.00	3.40	8.80	12.5307	13.2338
4.00	3.50	8.90	12.5307	13.3620
4.00	3.55	8.95	12.5307	13.4262
4.00	3.65	9.05	12.5307	13.5545

[1]The values for BF and PR used in this table are approximately distributed around the 1987 plant means of 3.561 and 3.240, respectively.

[2]The traditional milk pricing plan pays a base price (we use the average of values reported for the months in 1987) for one cwt. of milk at 3.5% butterfat test, and adds the differential to or subtracts it from this figure for each 0.1% difference above or below 3.5%.

[3]In each row of the table, the proposed price is calculated according to equation (5) in the text, with $IP_{BF} = 2.0208$, $IP_{PR} = 1.1399$, and $IP_{MSOL} = .1462$ (the 1987 average values). The totals of the last two columns are slightly different because of the approximation error introduced in our calculation of PREV, the total producer revenue.

Appendix: Mixed Estimation and Ridge Regression

Suppose we have the classical linear regression model

$$y = X\beta + \epsilon, \tag{A.1}$$

where y is n×1, X is n×k, β is k×1, and ϵ is n×1 with $E(\epsilon) = 0$, $var(\epsilon) = \Sigma_1$, and Σ_1 is a positive definite n×n matrix. Suppose, further, that there exist $p \geq 1$ linear prior restrictions which the researcher is willing to impose on the model. These are given by

$$c = R\beta + \xi, \tag{A.2}$$

with $E(\xi) = 0$ and $var(\xi) = \Sigma_2$, a positive definite p×p matrix. R is a known constant p×k matrix of rank $p < k$, c is a vector of specified constants, and ϵ is independent of ξ. The researcher also specifies Σ_2.

Combining (A.1) and (A.2), we have

$$\begin{bmatrix} y \\ c \end{bmatrix} = \begin{bmatrix} X \\ R \end{bmatrix} \beta + \begin{bmatrix} \epsilon \\ \xi \end{bmatrix}, \tag{A.3}$$

where the variance of the error term is given by

$$\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}. \tag{A.4}$$

In the event that Σ is known, generalized least squares may be implemented to calculate the mixed estimator as

$$\hat{\beta}_{ME} = \left(X^T\Sigma_1^{-1}X + R^T\Sigma_2^{-1}R\right)^{-1}\left(X^T\Sigma_1^{-1}y + R^T\Sigma_2^{-1}c\right). \tag{A.5}$$

Moreover, if one specifies $\Sigma_1 = s^2I_n$, then (A.5) becomes

$$\hat{\beta}_{ME} = \left(X^TX + s^2R^T\Sigma_2^{-1}R\right)^{-1}\left(X^Ty + s^2R^T\Sigma_2^{-1}c\right). \tag{A.6}$$

Some computer statistical packages, including the time-series package RATS, perform mixed estimation by pre-coded routines. Here, it is shown that (A.6) is easily manipulated in such a way that it may be estimated simply using matrix-based languages such as GAUSS or MATLAB. Denote by D the p×p matrix $D = s\Sigma_2^{-1/2}$. We show that (A.6) is equivalent to the OLS estimator of the model

$$\begin{bmatrix} y \\ Dc \end{bmatrix} = \begin{bmatrix} X \\ DR \end{bmatrix} \beta + \begin{bmatrix} \epsilon \\ D\xi \end{bmatrix}. \tag{A.7}$$

Let (A.7) be given by the expression $y^* = X^*\beta + \zeta$, where $\zeta^T = (\epsilon^T \; \xi^T D^T)$, and where X^* and y^* denote the obvious augmented matrices. Let the OLS estimator of (A.7) be denoted β^*. If we assume that $\Sigma_1 = s^2 I_n$, then the covariance matrix of the error term ζ may be shown to satisfy the restrictions of the ordinary linear model. It follows that

$$\text{var}(\zeta) = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & D^T\Sigma_2 D \end{bmatrix} = s^2 \begin{bmatrix} I_n & 0 \\ 0 & I_p \end{bmatrix}, \tag{A.8}$$

which clearly satisfies the assumptions required for the application of the Gauss-Markov theorem and the method of least squares. The estimator β^* is calculated from (A.7) as

$$\begin{aligned} \beta^* &= (X^{*T}X^*)^{-1}X^{*T}y^* \\ &= \left(X^TX + R^TD^TDR \right)^{-1} \left(X^Ty + R^TD^TDc \right) \\ &= \left(X^TX + s^2R^T\Sigma_2^{-1}R \right)^{-1} \left(X^Ty + s^2R^T\Sigma_2^{-1}c \right) = \hat{\beta}_{ME}. \end{aligned}$$

Usually, the prior information will not include restrictions on the covariances of the β_i's. That is, Σ_2 will usually be diagonal, so that D is easily calculated. In any event, while Σ_2 is positive definite, mixed estimation may be applied by adding the p-dimensional vector Dc to the dependent variable column vector and the p×k matrix DR to the design matrix as in (A.7) before performing OLS.

Ridge regression (Hoerl and Kennard, 1970) is a special case of mixed estimation. The ridge estimator is

$$\hat{\beta}_{RR} = \left(X^T X + dI_n\right)^{-1} X^T y, \qquad (A.9)$$

where $d \geq 0$ is the ridge parameter. While one may specify different weights on various of the β_i restrictions using mixed estimation, ridge regression imposes identical weights. What's more, restrictions of the kind used in the paper, where we specified

$$R = \left[\, 0\ 1\ 1\ 0\ 0\ 0 \,\right],$$

are impossible using ridge regression. Only individual parameter values may be restricted. If we take $\Sigma_1 = \sigma^2 I_n$ and $\Sigma_2 = \rho^2 I_p$, and if we let $c = 0$, then (A.5) is easily shown to be the ridge estimator with $d = \sigma^2/\rho^2$.

The stochastic prior ridge (SPR) estimator of Fomby and Johnson (see also Burt, Frank and Beattie) is given by

$$\hat{\beta}_{SPR} = \left(X^T X + dI_n\right)^{-1} \left(X^T y + dB \right), \qquad (A.10)$$

where B is a k-dimensional vector of random variables which are specified as constants by the researcher. This estimator lies between the mixed and ridge estimators in some sense. The ridge estimator biases the OLS estimate of β uniformly (that is, in each component) toward zero. The SPR estimator of Fomby and Johnson avoids this limitation. Using the SPR approach, each β_i may be biased toward a separate (possibly nonzero) real number. However, the SPR limits the researcher to restricting only individual parameters. As with the ridge estimator, the analysis of this paper, in which we imposed a restriction on the sum of two parameters, could not have been carried out using the stochastic prior ridge technique.

In addition, the SPR estimator is a special case of the mixed estimator

(A.5). To see this, suppose that $\Sigma_1 = \sigma^2 I_n$. Then if $\Sigma_2 = \rho^2 I_p$, we have $\hat{\beta}_{ME} = \hat{\beta}_{SPR}$ whenever

1.) $\sigma^2 R^T \Sigma_2^{-1} R = dI_n$ and

2.) $\sigma^2 R^T \Sigma_2^{-1} c = dB$.

But these hold whenever $d = \sigma^2/\rho^2$, $R = I_p$, and $c = B$.

These two features, the nature of the prior information provided and the degree of limitation on the form of the parameter restrictions, favor the mixed estimator quite strongly over both the ridge and the stochastic prior ridge estimators for confronting an ill-conditioned design matrix in linear regression analysis.

REFERENCES

Belsley, D.A., E. Kuh, and R.E. Welch. *Regression Diagnostics*. New York, NY: John Wiley and Sons, 1980.

Burt, O.R., M.D. Frank, and B.R. Beattie. "Prior Information and Heuristic Ridge Regression for Production Function Estimation." *Western J. Agr. Econ*. 12(1987):135-143.

Ernstrom, C.A. "End Product Pricing of Milk." Mimeo, Utah State University, 1983.

Fomby, T.B., and S.R. Johnson. "MSE Evaluation of Ridge Estimators Based on Stochastic Prior Information." *Commum. Stat*. 6(1977):1245-58.

Froker, R.K. and C.M. Hardin. "Paying Producers for Fat and Solids-Not-Fat in Milk." Wisconsin Agr. Exp. Sta. Res. Bull. 143 (1942).

Graf, T.F. "Economic Rationale for Using Both Non-Fat and Fat Values in Pricing Milk." Univ. of Wisconsin College of Agr. Research Report 129, 1977.

Hillers, J.K., V.H. Nielson, A.E. Freeman, J. Dommerholt, and R.E. Deiter. "Value of Fat and Protein in Producer Milk." *J. Dairy Sci*. 63(1980): 322-327.

Hoerl, A.E. and R.W. Kennard. "Ridge Regression: Biased Estimation for Nonorthogonal Problems." *Technometrics*. 12(1970):55-68.

Ladd, George W. and John R. Dunn. "Estimating Values of Milk Components to a Dairy Manufacturer." *J. Dairy Sci*. 62(1979):1705-1712.

Lelievre, J., O.J. Freese, and J. Gilles. "Prediction of Cheddar Cheese Yield." *New Zealand J. of Dairy Science*. 18(1983):169-72.

Perrin, R.K. "The Impact of Component Pricing of Soybeans and Milk." *Amer. J. Agr. Econ*. 62(1980):445-55.

St-Pierre, Normand R. and Grant M. Scobie. "The Component Pricing of Milk Revisited." *Amer. J. Agr. Econ*. 69(1987):697-700.

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

Theil, H. and A.S. Goldberger. "On Pure and Mixed Statistical Estimation in Economics." *Int. Econ. Review*. 2(1961):65-78.

Van Slyke, Lucius L. and Walter V. Price. *Cheese*. New York: Orange Judd Publishing Co., Inc., 1952.