# Pseudo Data as a Teaching Tool: Application to the Translog, Multiproduct Profit Function

**Thomas W. Hertel and Lance McKinzie**

This paper argues that the use of "laboratory" data sets can add substantially to the teaching of production economics at the graduate level. Optimal experimental designs for generating pseudo data from a process model are discussed. These are shown to depend on the functional form to be estimated. We choose the translog form for our multiproduct profit function and compare alternative approaches to estimation, using pseudo data from a farm-level linear programming model. Particular restrictions on this profit function are also considered. Finally, aggregation of output prices is shown to alter substantially input price elasticities of demand.

*Key words:* multiple products, profit function, pseudo data, teaching, translog.

Recent developments in duality theory and the concept of flexible functional forms have led to a resurgence of interest in production economics. Whether the topic is factor substitution, income distribution, technical change, economies of scale, or any of the other traditional problems in production theory, the current literature draws heavily on these new methods. The use of duality in production theory dates back to Shephard's 1953 work. Applications involving flexible forms have been available for almost two decades (e.g., Diewert 1969). However, widespread incorporation of these methods into graduate production economics curricula is much more recent. A valuable set of papers on this topic is provided in the December 1982 issue of this journal, titled: "Relevance of Duality Theory to the Practicing Agricultural Economist."

The challenge in teaching this newer material to students in production economics lies not only in conveying the theoretical concepts but also in teaching their responsible application. When is it appropriate to apply these more sophisticated methods, and what can go wrong when they are applied? The traditional approach to this type of teaching challenge has been to give students an empirical problem to work with. In the case at hand, this would involve giving them a data set with which to estimate, e.g., a profit function, which could in turn be interpreted and perhaps criticized. A second-best alternative might be to assign a set of empircal articles to be read and evaluated.

Unfortunately, many of the data sets in use are not "well-behaved," i.e., estimation of a dual function using these observations does not result in a set of parameters which satisfies the required neoclassical restrictions. This problem is particularly severe in the multiproduct setting (e.g., Shumway). Strictly speaking, the duality results do not apply to these ill-behaved functions, and we are left with something which cannot be readily interpreted. This problem may, or may not, be acknowledged by the authors of journal articles. In most cases the amount of information provided is insufficient for readers to check for themselves whether or not the function is well-behaved.

When a profit function is found to be ill-behaved, explanations generally turn to problems with the data set (Pope; Lopez 1982;

---

Thomas W. Hertel and Lance McKinzie are, respectively, an assistant professor and a research associate in the Department of Agricultural Economics, Purdue University.

The authors would like to thank Dick Boisvert and Dale Squires for valuable conversations on this topic. Participants in the Advanced Production Economics Seminar at Purdue University also contributed useful comments. Virgil Anderson and Charles Oellerman provided insights into the problem of experimental design. Paul Preckel assisted with computational algorithms.

*Western Journal of Agricultural Economics*, 11(1): 19–30

Young). Poor quality data and excessive aggregation (over commodities and/or firms) are commonly cited sources of difficulty. Another potential pitfall is that the underlying behavioral axioms (e.g., profit maximization or cost minimization) may not be satisfied. In some cases the latter are posed as testable hypotheses (Appelbaum), although they are generally maintained. Studies which attribute poor results to one of these causes can serve a valuable purpose in graduate courses, illustrating the point that not all research succeeds. However, they do not provide students with an adequate feeling for what the methods are good for.

Consistently attributing bad results to poor data encourages a certain cynicism and sometimes sloppiness on the part of students who resign themselves to the fact that "this never works out in practice anyway." What is needed is a "laboratory" data set which permits teachers to abstract from data deficiencies, thus enabling students to focus attention on the method—how it is used, and what its strengths and weaknesses are. Of course, such a data set could also be selectively "disrupted" (e.g., via inappropriate aggregation) to illustrate the potential damage which can result.

There are several properties which a laboratory data set should satisfy. These include: (*a*) The underlying technology is sufficiently well understood to permit formulation of preliminary hypotheses. These should in turn facilitate interpretation and discussion of the estimated model. (*b*) The behavioral axioms guiding producers are known and conform with neoclassical postulates. (*c*) The observed data are accurate and not aggregated. (*d*) Price variability is sufficient to permit measurement of all distinct substitution effects.

This paper reports the development of a "pseudo" data set which meets (*a*)–(*d*) and illustrates its use in the estimation of a flexible, multiproduct profit function. While some of our results are informative in their own right, their greatest value has been in the teaching of a graduate-level production economics course. The second section of the paper summarizes results for the multiproduct translog profit function to be estimated. The linear programming model from which pseudo data is generated is discussed next. The fourth section addresses the question of optimal experimental designs for generating pseudo data. Estimation and interpretation of the translog multiproduct profit function are the subjects of the next section. The final section provides a summary and some conclusions.

## The Translog Multiproduct Profit Function

We have chosen to illustrate the use of pseudo data as a teaching tool by applying this concept to the case of the translog, multiproduct profit function. As noted above, the multiproduct case is obviously more general than its single-product counterpart. Thus, interpretation is more difficult but also more interesting. The translog functional form has been selected on the basis of its popularity. It is the most widely used of the "second-order flexible" functional forms. In order to facilitate later discussion, this section provides a convenient summary of relevant results pertaining to the translog, multiproduct profit function. Following Blackorby, Primont, and Russell, we write this as

$$\ln \Pi = \mu_0 + \Sigma_i \mu_i \ln P_i + \tfrac{1}{2}\Sigma_i \Sigma_j \mu_{ij} \ln P_i \ln P_j \tag{1}$$

where

$$\mu_o = \bar{\mu}_o - \Sigma_i \bar{\mu}_i \ln \bar{P}_i + \tfrac{1}{2}\Sigma_i \Sigma_j \bar{\mu}_{ij} \ln \bar{P}_i \ln \bar{P}_j, \tag{2}$$

$$\mu_i = \bar{\mu}_i - \Sigma_j \bar{\mu}_{ij} \ln \bar{P}_j, \tag{3}$$

$$\mu_{ij} = \bar{\mu}_{ij}\,, \tag{4}$$

and $\bar{P}$ is the netput price vector around which the true profit function is being approximated. It contains both output and input prices. This serves to simplify the notation in the text of this paper considerably. An alternative expression for this multiproduct profit function is provided at the bottom of table 1. There, output prices ($p_i$: $i = 1, \ldots, m$) are distinguished from input prices ($w_k$: $k = 1, \ldots, n$) thus introducing more extensive notational requirements.

The parameters $\bar{\mu}_0$, $\bar{\mu}_i$, and $\bar{\mu}_{ij}$ in (2)–(4) are the zero-, first-, and second-order derivatives of the true function at the point of approximation. In the translog case, setting $\bar{P} = \{1\}$ results in $\mu_0 = \bar{\mu}_0$, $\mu_i = \bar{\mu}_i$, so that the estimated parameters may be associated directly with the underlying "true" function. Symmetry ($\mu_{ij} = \mu_{ji}$) and linear homogeneity ($\Sigma_i \mu = 1$, $\Sigma_j \mu_{ij} = \Sigma_i \mu_{ij} = 0$) restrictions may also be imposed on (1). Monotonicity and convexity cannot be easily imposed by a set of linear restrictions on the parameters (Lau 1978). Thus, it has become conventional to estimate the profit function under the linear restrictions, thereafter checking whether or not the remaining restrictions are violated.

### *Deriving Output Supply and Input Demand Elasticities*

Gross elasticities are written in terms of second-order coefficients and predicted shares (Binswanger). At the point of approximation, the latter are equal to the first-order coefficients. Thus, at the unit price vector we have

$$(5) \qquad z_{ij} = \frac{\bar{\mu}_{ij}}{\bar{\mu}_i} + \bar{\mu}_j, \qquad i \neq j,$$

and

$$z_{ii} = \frac{\bar{\mu}_{ii}}{\bar{\mu}_i} + \bar{\mu}_i - 1.$$

When $i$ denotes an output, than $z_{ij}$ represents a gross cross-price elasticity of supply with respect to output or input prices and $z_{ii}$ is the gross own-price elasticity of supply. Similarly, when $i$ denotes an input, $z_{ij}$ and $z_{ii}$ represent cross- and own-price elasticities of demand. Compensated elasticities for the translog case can be derived, based on the Sakai/Lopez decomposition results (Hertel). They are summarized for the base price vector as follows ($i$, $j \in m$ outputs; $k$, $h \in n$ inputs).

Compensated output supply elasticities:

$$(6) \qquad \underset{(m\times m)}{\{\nu_{ij}\}} = \underset{(m\times m)}{\{z_{ij}\}} - \underset{(m\times n)}{\{z_{ik}\}}\underset{(n\times n)}{\{z_{kh}\}^{-1}}\underset{(n\times m)}{\{z_{ki}\}}.$$

Compensated input demand elasticities:

$$(7) \qquad \underset{(n\times n)}{\{\nu_{kh}\}} = \underset{(n\times n)}{\{z_{kh}\}} - \underset{(n\times m)}{\{z_{ki}\}'}\underset{(m\times m)}{\{z_{ij}\}^{-1}}\underset{(m\times n)}{\{z_{ik}\}}.$$

### *Testing for Nonjointness and Separability*

It is common to test whether or not additional restrictions on the profit function are violated. In the case of the multiproduct firm, both output separability and nonjointness are of particular interest. Separability of outputs from inputs implies that the ratio of optimal output supplies does not vary with changes in input prices. This simplifies the problem of modeling a multiproduct firm by permitting outputs to be aggregated. By contrast, the nonjointness (in inputs) restriction simplifies matters by legitimizing the use of single-commodity production (profit) functions. The implication is that decisions about the production of any one commodity are independent of similar decisions about other outputs.

It can be shown (Blackorby, Primont, and Russell, p. 298) that the output separability condition is equivalent to the following restriction ($i, j \in$ outputs, $h \in$ inputs, and $l \in$ inputs and outputs):

$$(8) \qquad \mu_j\mu_{ih} - \mu_i\mu_{jh} + \sum_l (\mu_{jl}\mu_{ih} - \mu_{il}\mu_{jh})\ln(p_l) = 0.$$

This equality holds for all $P$ if and only if

$$(9) \qquad \mu_j\mu_{ih} - \mu_i\mu_{jh} = 0,$$

and

$$(10) \qquad (\mu_{jl}\mu_{ih} - \mu_{il}\mu_{jh}) = 0 \text{ for all } l.$$

At this point the question arises: Are we interested in testing for global separability of the estimated profit function, or are we interested in testing for separability in the true but unknown profit function which is being approximated? As Blackorby, Primont, and Russell; and Denny and Fuss have pointed out, this is an important distinction. At the point $\bar{P} = 1$, with $\ln(p_l) = 0$, (8) simplifies and (9) becomes a sufficient condition for separability. The resulting restrictions are summarized at the top of table 1, using parametric notation which distinguishes between input and output prices.

Global separability requires that (9) and (10) hold simultaneously. Berndt and Christensen have identified two alternative sets of sufficient restrictions in the translog case. The first, termed "linear separability," involves setting all parameters associated with input-output price interactions equal to zero (see table 1). Thus, not only are outputs separable from inputs, the reverse is true as well, giving rise to the term "input-output separability." The nonlinear restrictions for global separability are less restrictive, as inputs are no longer constrained to be separable from outputs. As noted in table 1, both of these global separability restrictions on the translog have strong implications for functional form.

Lau (1978) has shown that a necessary and sufficient condition for nonjointness in inputs is that all cross-derivatives of the profit function with respect to product prices are equal to zero. In the case of the translog profit function, this restriction cannot be globally ap-

**Table 1. Summary of Coefficients and Particular Restrictions**

| Restrictions | Comments |
|---|---|
| <u>Output separability restrictions</u> | |
| Approximate: $\alpha_j\gamma_{ih} - \alpha_i\gamma_{jh} = 0$, for all $i, j \in$ outputs, $h \in$ inputs. | Holds only at $\bar{P} = 1$. |
| Global linear: $\gamma_{ih} = \gamma_{hi} = 0$, for all $i \in$ outputs, $h \in$ inputs. | Translog profit function reduces to a Cobb-Douglas function of translog aggregates (Denny and Fuss). |
| Global, nonlinear: $\alpha_i/\alpha_j = \alpha_{im}/\alpha_{jm} = \gamma_{ih}/\gamma_{jh}$, for all $m, i, j \in$ outputs, $h \in$ inputs. | Output price aggregate is Cobb-Douglas in form (Denny and Fuss). |
| <u>Approximate nonjointness</u> | |
| $\alpha_{ij} = -\alpha_i\alpha_j$, $i \neq j \in$ outputs. | Holds only at $\bar{P} = 1$, global nonjointness is incompatible with the translog profit function. |

Note: Notation in this table applies to the following translog profit functions, where $p_{i,j}$ refer to output prices and $w_{k,h}$ refer to input prices:

$$\ln \Pi = \alpha_0 + \Sigma_i\alpha_i \ln p_i + \Sigma_k\beta_k \ln w_k + \tfrac{1}{2}\Sigma_i\Sigma_j\alpha_{ij} \ln p_i \ln p_j + \tfrac{1}{2}\Sigma_k\Sigma_h\beta_{kh} \ln w_k \ln w_h + \tfrac{1}{2}\Sigma_i\Sigma_h\gamma_{ih} \ln p_i \ln w_h + \tfrac{1}{2}\Sigma_k\Sigma_j\gamma_{kj} \ln w_k \ln p_j.$$

plied. Thus, the translog is in some sense "nonjoint inflexible." However, it is possible to formulate a test for approximate nonjointness in inputs at the base point (Denny and Pinto). This is provided in table 1. Rejection of this property at $\bar{P} = 1$ does not imply anything about nonjointness at other points on the profit surface. As such, it is a weaker test than would be provided by functional forms where a global restriction is available (e.g., the normalized quadratic profit function).

## The Vehicle: A Process Model

The process model used to generate pseudo data is a modified version of the Purdue Crop Budget Model, B-9 (McKinzie, Hertel, and Preckel). B-9 is among the most extensively validated of all process models, having been used daily by extension and research staff, graduate and undergraduate students, as well as by thousands of Midwest farmers over the course of its fifteen-year evolution. It is a linear programming formulation of a profit-maximizing farm firm. The formulation utilizes highly detailed information including the farm's machinery working rates, available time for working in the field during different periods of the production year, and cultivation practices.

Timing of production activities is given particular attention in the B-9 model. Expected crop yields generally are acknowledged to decline as planting (and harvesting) of the crop are delayed. However, it is not economical to maintain the necessary machinery set to plant (harvest) all of the crop at one time. The B-9 model captures tradeoffs between the cost of larger, more expensive machinery sets and the benefits associated with improved yields due to timeliness of planting and harvesting. The latter also serves to promote diversification among crop outputs. While corn is often the most profitable crop to be planted during late April and early May, soybeans may be the preferred alternative in late May. This occurs because soybean yields decline at a slower percentage rate than do corn yields as planting is delayed.

The first major modification of B-9 involved permitting all inputs, with the exception of land, to be purchased or rented at an exogenous price. (The quantity of land was held fixed in order to maintain a bounded solution.) Variable inputs now include labor, machinery, combine, storage, drying, fertilizer, chemicals, and other inputs. Each of these inputs is available in continuous quantities.[1] Several minor

[1] Abstracting from the discrete nature in which machinery must be employed will not be restrictive as long as the model is used to make inferences about the entire sector, as opposed to individual farms. An alternative justification for this assumption would be an active rental market. There are eight different machinery choice variables—four sizes of combines and four sizes of machinery complements which include all other machinery. The four sizes correspond roughly to 200-, 600-, 800-, and 1,200-acre operations. The different working rates for each machinery complement and combine are taken from Edelman. A typical solution for machinery might be .3 units of machinery complement no. 2 plus .5 units of machinery complement no. 3. These solution values are weighted by the respective annualized costs for each machinery set (Leatham and Baker) to arrive at a dollar cost for machinery. The combine decision is handled in an identical manner.

aspects of B-9 were dropped (e.g., silage) to simplify things. Substantive changes in the structure and parameters of B-9 were not necessary.

The second major modification of B-9 involved a more complete treatment of corn-soybean rotation. Typically, both input requirements and expected yields depend upon which crop was grown on the land in the previous year. There are significant economies from rotating corn and soybeans. Fertilizer and pesticide costs rise for corn grown continuously on the same land. Yields decline for both continuous corn and continuous soybean crops. The effect of including these complementarities from crop rotation is to give the product transformation curve for corn and soybeans more curvature in the region of equal acreages. Thus, holding other things constant, a greater change in relative prices is required to achieve a given amount of substitution between these crops. Rotation corn-soybeans was incorporated as an additional crop alternative with greater yields and lower fertilizer requirements compared to either crop on a continuous basis (*Farm Planning and Financial Management*).

In sum, the modified model provides an excellent vehicle with which to generate pseudo data for use in teaching production economics. Benefits from timeliness of operation and crop rotation make the multiple output approach particularly interesting. The basic model is quite detailed and extensively validated. These features facilitate the formulation of a priori hypotheses as well as the interpretation of the resulting profit function.

## Generating Pseudo Data

The concept of pseudo data was introduced into the economics literature by Griffin (1977a, b, 1978) as a means of summarizing the information embodied in industry process models. The resulting cost or profit functions may, in turn, be employed to summarize an individual sector's price responsiveness in large econometric models. Griffin's approach to generating observations begins with the selection of a base point consisting of the unit price vector. This coincides conveniently with the point at which the parameters of his estimated translog function may be directly related to the derivatives of the underlying "true" function. Griffin then proceeds to generate sample points

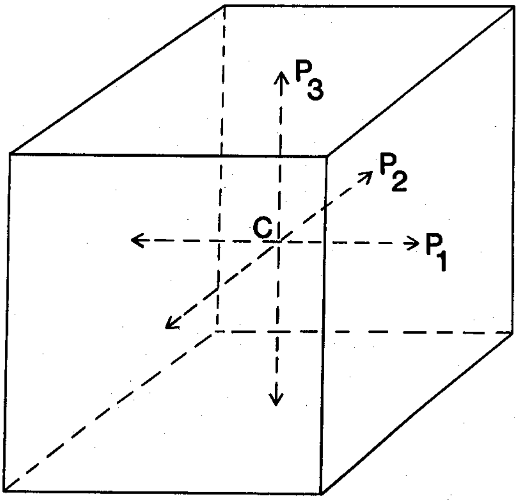


**Figure 1. Experimental designs for three factors**

by holding all but one price constant (at base values) while varying the one remaining price over a predetermined range. This is illustrated for the three-price case in figure 1. The base case: $C = (1, 1, 1)$ falls in the interior (in this case the center) of the sample space, which is represented here by a cube. Griffin's procedure results in observations falling along the $P_1$, $P_2$, and $P_3$ axes. These axes intersect at the base case ($C$), which may be viewed as a translation of the origin.

Critics of Griffin's method (Maddala and Roberts) point out that the estimated coefficients may be quite sensitive to sample design and the number of basis changes which result from the price configurations chosen. Griffin himself (1982) has noted the sensitivity of his results to the frequency and range over which sample points are selected. It is somewhat surprising, then, that none of these authors make reference to the extensive literature on experimental design. This literature draws an explicit link between choice of design and the form of the regression model. Thus, if the objective is to estimate a second-order Taylor series approximation, such as the translog, it is important to choose a compatible experimental design. For ease of reference, consider the following response surface with three factors ($x_1$, $x_2$, $x_3$) determining the level of the dependent variable ($y$):

$$(11)\quad y = \mu_0 + \mu_1 x_1 + \mu_2 x_2 + \mu_3 x_3 + \mu_{11} x_1^2 + \mu_{22} x_2^2 + \mu_{33} x_3^2 + \mu_{12} x_1 x_2 + \mu_{13} x_1 x_3 + \mu_{23} x_2 x_3.$$

By defining $y = \ln \Pi$ and $x_i = \ln p_i$, this becomes a translog profit function.

An important question in designing the experiment is: Over how many levels must each factor ($x_i = \ln p_i$) be varied (at a minimum)? Two levels for each price is sufficient to estimate $\mu_0$, the linear terms and the interaction effects ($\mu_{12}$, $\mu_{13}$, $\mu_{23}$) (Anderson and McLean, p. 353). A two-level factorial for three factors (prices) generates $2^3 = 8$ observations. If each price takes on values of, e.g., 0.75 and 1.25, the sample points could be represented by the corners of the cube in figure 1 with the center ($C$) having the coordinates (1, 1, 1).

In order to capture the quadratic effects ($\mu_{11}$, $\mu_{22}$, $\mu_{33}$) we need to go to three levels for each price (Anderson and McLean, p. 323). A full three-level factorial generates $3^3 = 27$ observations. In terms of figure 1, this would involve sampling the midpoint of each segment. That is, each price could now take on the values 0.75, 1.0, or 1.25. The problem with a full three-level factorial is that, as the number of factors increases, the number of observations required becomes rapidly unmanageable. Thus, for the eleven-price profit function considered here, $3^{11} = 177{,}147$ data points result.

Fortunately, it is not necessary to carry out a full three-level factorial in order to estimate all of the coefficients in the regression model (11). A composite design will permit measurement of both interaction and quadratic effects. This design consists of three groups of observations.

(*i*) A two-level factorial (the 8 corners of the box).

(*ii*) Points at the extreme of each factor, while at the center of the others (the center point of each of the cube's six faces, i.e., $P_i = 0.75$, 1.0, and 1.25 with $P_j = 1.0, j \neq i$).

(*iii*) The center point itself (point $C$).

Together (*i*)–(*iii*) total $8 + 6 + 1 = 15$ observations.

The required number of sample points for a composite design rises much more slowly, as the number of factors increases, than is the case with the three-level factorial. Thus, for the eleven-price case "only" $2^{11} + (2 \times 11) + 1 = 2{,}071$ sample points are required. This number may be further reduced by utilizing a fractional two-level factorial in the composite design. In this case a $\frac{1}{16}$th fractional factorial is employed.[2] This results in $(2^{11}/16 = 128$ "corner" points. Adding the 22 factor extrema (*ii*) and the center point (*iii*) yields a sample of 151 observations.

It is interesting to note that the directions associated with (*ii*) and (*iii*) (Griffin's design) are sufficient for estimating both quadratic and interaction terms if, instead of estimating the profit function directly, we estimate the associated system of share equations which is linear in prices. In effect, use of Hotelling's lemma permits a further reduction in sample size. However, in this paper we are interested in being able to estimate the profit function directly. Therefore, the sample points associated with (*i*) are also required.

Two further design considerations remain to be discussed. The first is the issue of orthogonality. In order to make the proposed composite design orthogonal, the extrema in (*ii*) must be moved off of the surface of the cube. The distance from the center to these points becomes a function of the number of treatment combinations. This changes the form of the polynomial in equation (11) (Anderson and McLean, pp. 356–59), such that it no longer corresponds to the translog formulation. As a result, the design employed here will not be orthogonal.

A final design consideration relates to the appropriate distance between sample points. Griffin (1982) found his results to be quite sensitive to the spacing and frequency of his pseudo data points. Unfortunately, there is little one can say, in general, other than that the changes should be large enough to cause basis movements in the underlying process model.

## Estimation and Interpretation of Results

When applied to the pseudo data set outlined in the previous section, the translog profit function in equation (1) will not hold exactly. However, unlike conventional econometric applications where there is a multiplicity of reasons for an error term to exist, there is only one in this case—namely, approximation error. Thus, there is no reason to believe these errors

[2] Anderson recommended the $\frac{1}{16}$th factorial as the minimum necessary to measure the required two-way interactions (personal communication). The reference for such designs is a 1957 monograph (Statistical Engineering Laboratory). Plan 16.11.64 (p. 46) supplied the design for this research.

will be normally distributed (Maddala and Roberts). However, since we already know the underlying model, statistical inference per se is not our objective; and so we will avoid formal hypothesis testing in this section. Rather, emphasis will be placed on point estimates and their interpretation.

Because of the collinearity of most "real world" price series, it is rare for researchers to estimate the profit function as a single equation. Instead, the system of $(m + n - 1)$ share equations (sometimes including the profit function) is estimated. The associated error terms are conventionally attributed to errors in profit maximization and are likely to be contemporaneously correlated. This is used to justify a seemingly unrelated regression (SUR) technique. However, all of the regressors are identical in each share equation, and the SUR estimator would collapse to OLS were it not for the cross-equation constraints. Thus, it is really the latter which justifies estimation of a system of share equations.

Because one of the share equations must be dropped, it is often recommended that the researcher iterate over the covariance matrix to assure invariance of the estimates to the choice of equations (Berndt and Christensen). However, there are no known statistical gains with this procedure. In fact, in their monte carlo experiments, Kmenta and Gilbert find that the simple SUR estimator performs equally as well as the more costly iterative procedures, even when the model is substantially misspecified.

Using our pseudo data set, we are able to explore all of these approaches to estimation of the eleven-price, translog profit function. We begin by estimating the profit function directly. The share equation system is then estimated, both with and without iteration over the covariance matrix.[3] Estimates from both of these models are compared. This is followed by a more detailed investigation of the multiproduct technology implicit in the process model. The first step is to impose approximate nonjointness. This is followed by the imposition of successively more restrictive separability conditions. The latter determine the validity of a single output profit function, the estimates of which are presented at the end of the section.

[3] Attempts to estimate a system comprised of the profit function as well as the share equations were unsuccessful, as convergence was not obtained.

### *Comparing Alternative Approaches to Estimation*

Direct estimation of the profit function, with symmetry and positive linear homogeneity imposed, involves ten linear terms, ten quadratic terms, and forty-five interaction terms. Since the experimental design has not been orthogonalized, collinearity is a potential problem in the ordinary least squares estimation of this equation containing sixty-six variables, although there is sufficient independent variability to estimate the model.[4] The "fit" of the equation resulted in a high $R^2$: 0.988. This is not surprising, based on the experience of earlier authors with pseudo data sets (Maddala and Roberts; Griffin). Only nine of the fifty-five second-order coefficients exceeded their associated standard error by a factor of two or more.

Application of Hotelling's lemma to the profit function results in ten independent share equations which were normalized on "other inputs":

$$\frac{\partial \ln \pi}{\partial \ln P_i} = S_i = \mu_i + \Sigma_j \mu_{ij} P_j, \quad i = 1, \ldots, 10. \tag{12}$$

Here $S_i$ is positive for outputs and negative for inputs. These shares sum to one by definition. In addition to reducing collinearity, this system approach has the advantage of creating $(11 \times 151)$ "observations" without increasing the number of parameters to be estimated. As a result, system estimation yields forty-five of fifty-five second-order coefficients which exceed their standard error by a factor of two or more. As noted above, iteration over the covariance matrix has the advantage of providing estimates which are invariant to the missing share equation. In the case of the particular data set and model utilized here, iteration over the covariance matrix produced no change in most coefficients. Only one estimate changed by more than one percent (1.86%).

A final alternative is to estimate the profit function and share equations jointly. This

[4] In order to make the proposed composite design orthogonal, the distance from the center point to the extrema in part (*ii*) of the design must be adjusted. It becomes a function of the number of treatment combinations. This in turn changes the form of the corresponding regression model (1) (Anderson and McLean, pp. 356–59), such that it no longer corresponds to the translog formulation.

**Table 2. Actual and Predicted Profit Shares for the Base Case: Two Different Estimators**

| | | Predicted Values | |
|---|---|---|---|
| | Actual Values | Profit Function | Share Equations |
| Corn ($\alpha_1$) | .997 | .480 | .962 |
| Soybeans ($\alpha_2$) | .786 | .893 | .834 |
| Wheat ($\alpha_3$) | .111 | .480 | .147 |
| Labor ($\beta_1$) | −.172 | −.168 | −.180 |
| Machinery ($\beta_2$) | −.02 | −.016 | −.024 |
| Combine ($\beta_3$) | −.033 | −.034 | −.041 |
| Drying ($\beta_4$) | −.024 | −.014 | −.025 |
| Storage ($\beta_5$) | −.056 | −.045 | −.054 |
| Fertilizer ($\beta_6$) | −.263 | −.252 | −.269 |
| Chemicals ($\beta_7$) | −.097 | −.099 | −.109 |
| Other ($\beta_8$) | −.229 | −.225 | −.241 |

brings the maximum amount of information to bear on the estimation problem. However, it is also a very difficult and expensive problem when the number of prices is large, as in the case at hand. As a result we were not able to obtain convergent estimates of the joint system for this particular model.

Table 2 provides a comparison of results from the estimated profit function and share equation models. The first column lists actual profit shares for each of the three outputs and eight inputs at the point of approximation ($\bar{P} = 1$).[5] These are derived by dividing net revenues (or negative costs in the case of inputs) by returns to land and management. Estimates of these actual profit shares are provided by the first-order terms in the translog profit function. The second and third columns of table 2 list these estimates from the profit function and share equation models, respectively. Note that output shares are positive and input shares negative, thus satisfying the monotonicity property (convexity will be discussed below). The most striking weakness of the profit function approach is its inability to make reasonable predictions of the output shares of corn and wheat in the base case. The OLS estimate of $\alpha_1$ is far too low, while that for $\alpha_3$ is far too high. Since elasticities in this model are a function of estimated shares, this problem is also likely to interfere with the profit function's ability to predict price effects. In short, the share equation approach, which predominates the literature, appears to be preferable in the case at hand.

*Interpretation of the Results*

Using formulas (5)–(7), gross and net elasticities were computed based on the share equation estimates of the translog profit function. These elasticities vary over the entire sample space and are derived as a function of estimated shares, so that standard errors could not be computed. We have chosen to evaluate them at the point of approximation.

Table 3 summarizes our estimates of the net elasticities in the base case. The input-compensated, output supply elasticities are generally less than one, with positive own-price effects and very large cross-price effects. Corn and wheat are found to be net complements in production. That is, an increase in the price of wheat leads to an increase in the optimal supply of corn, and vice versa. (Remember that land is being held constant throughout this analysis.) This can be explained by focusing on the major reason for crop diversification, namely, the timing of production activities. Consider what happens when wheat production increases in response to improved wheat prices while input availability is held constant. The shift of land into wheat lessens demands on labor and machinery during the spring and fall. With these resources less constrained, it is profitable to shift land from soybeans to the more input-intensive corn production. The latter are strong net substitutes, competing keenly for inputs during the planting and harvesting periods.

Output-constant, input demand elasticities are quite small, indicating that most of the "action" in this model comes from changes in output mix. While there are many different activities with which to produce, e.g., corn, they involve very similar input mixes. Thus, the own-price effects are quite inelastic. The only cross-price elasticity exceeding 0.10 (in absolute value) is the compensated demand elasticity for combines with respect to changes in the price of labor (0.22).

As expected, when all choice variables are permitted to adjust optimally, the firm's responses to price changes are more elastic. Thus,

[5] There is a contradiction inherent in treating these estimated coefficients as derivatives of the true underlying function at the point of approximation. This arises due to the nature of the remainder terms in a Taylor-series expansion, which increase as one moves away from the point of approximation. Unfortunately, regression attempts to make these errors "equally small" over the entire range of data. White has shown that least squares can produce only unbiased estimates of the Taylor-series coefficients if the estimated and true functions share the same form.

**Table 3. Net Elasticities**

| | Price | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Quantity | Corn | Soybeans | Wheat | Labor | Machinery | Combine | Drying | Storage | Fertilizer | Chemicals | Other Inputs |
| Corn | .5808 | −.6929 | .1116 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Soybeans | −.7988 | 1.0104 | −.2111 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Wheat | .7302 | −1.1983 | .4678 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Labor | 0 | 0 | 0 | −.1326 | −.0116 | .0507 | .0030 | .0034 | .0310 | .0120 | .0442 |
| Machinery | 0 | 0 | 0 | −.0862 | −.1182 | −.0083 | .0181 | .0156 | .0772 | .0271 | .0753 |
| Combine | 0 | 0 | 0 | .2213 | −.0049 | −.3201 | −.0006 | −.0096 | .0253 | .0304 | .0592 |
| Drying | 0 | 0 | 0 | .0218 | .0177 | −.0010 | −.1039 | .0164 | .0220 | .0284 | .0004 |
| Storage | 0 | 0 | 0 | .0113 | .0070 | −.0073 | .0075 | −.0295 | −.0149 | .0170 | .0086 |
| Fertilizer | 0 | 0 | 0 | .0207 | .0070 | .0039 | .0020 | −.0030 | −.0333 | .0090 | −.0061 |
| Chemicals | 0 | 0 | 0 | .0198 | .0060 | .0115 | .0065 | .0084 | .0222 | −.0425 | −.0319 |
| Other inputs | 0 | 0 | 0 | .0332 | .0076 | .0102 | .0000 | .0019 | −.0069 | −.0145 | −.0316 |

the gross elasticities shown in table 4 include an expansion effect in addition to the previous substitution effect. All of the outputs are gross substitutes, and the cross-price effects between corn and soybeans dominate the soybean own-price elasticity, indicating strong substitution possibilities in addition to the larger budget share for corn. The gross supply elasticity for wheat is very large, in part due to its relatively small share in the base case.

Gross input demand elasticities are generally quite close to their output-compensated counterparts in table 3. However, a few changes merit comment. The own-price elasticity of demand for chemicals is now positive, although not distinguishable from zero. The fact that this diagonal element of the Hessian carries the wrong sign indicates that the profit function cannot be convex in prices at the point of approximation. This is disturbing indeed, as we know that the underlying "true" profit function must be convex. However, closer examination of the laboratory data set indicates that chemical applications per acre are virtually unchanged across all solutions of the process model. Thus, the true value of that gross elasticity is zero. Since the estimated profit function is only an approximation, it is not surprising that the own-price elasticity carries an "insignificant" but wrong sign. In such a case it would seem that imposition of the convexity restriction (following Lau 1978) would be desirable, given the availability of appropriate software.

Another notable point in table 4 is the gross complementarity between fertilizer and drying inputs. This arises from the fact that they are both important inputs in the production of corn. An increase in the price of fertilizer causes a drop in the optimal supply of corn. This in turn dampens the demand for the complementary drying input, which is also intensively employed in corn production.

Note that all of the inputs are regressive against soybeans. That is, an increase in the price of any input results in an increase in the optimal supply of soybeans. Symmetrically, an increase in the price of soybeans results in a drop in the demand for any of the inputs. The opposite is true for corn and input demands. These results follow from two facts: (*a*) total land area is fixed, and (*b*) soybeans are relatively less input-intensive than corn.

### *Particular Restrictions*

In this section we evaluate the compatibility of the process model with the nonjointness and separability restrictions developed above. As noted above, the nature of the error term in pseudo data sets presents problems for classical statistical inference. Thus, rather than presenting formal test statistics, we have chosen simply to present system sum-of-squared residuals (based on FIML residuals) resulting from the restricted models. These are presented below.

Sum-of-Squared Residuals for Particular Restrictions

| Restriction | System SSR |
|---|---|
| Symmetry and positive linear homogeneity (maintained hypothesis) | 77.37 |
| Approximate separability | 80.35 |
| Global, nonlinear separability | 120.60 |
| Global, linear separability | 279.95 |
| Approximate nonjointness | 420.58 |
| Cobb-Douglas | 964.13 |

**Table 4. Gross Elasticities**

| | Price | | | | | | |
|---|---|---|---|---|---|---|---|
| Quantity | Corn | Soybeans | Wheat | Labor | Machinery | Combine | Drying |
| Corn | 1.5055 | −1.0772 | −.1196 | −.0369 | −.0087 | −.0145 | −.0525 |
| Soybeans | −1.2418 | .7895 | −.0562 | .0625 | .0110 | .0035 | .0426 |
| Wheat | −.7823 | −.3189 | 1.8754 | −.1526 | −.0291 | .0631 | .0261 |
| Labor | .1970 | −.2896 | .1246 | −.1427 | −.0149 | .0514 | −.0043 |
| Machinery | .3425 | −.3768 | .1761 | −.1105 | −.1237 | −.0061 | .0059 |
| Combine | .3379 | −.0716 | −.2252 | .2244 | −.0036 | −.3279 | −.0119 |
| Drying | 2.0268 | −1.4279 | −.1540 | −.0311 | .0058 | −.0198 | −.1745 |
| Storage | .6186 | −.6263 | −.1252 | .0255 | .0061 | −.0202 | −.0150 |
| Fertilizer | .5806 | −.6952 | .1624 | .0047 | .0008 | .0009 | −.0191 |
| Chemicals | .0837 | −.3064 | −.0113 | .0431 | .0070 | .0053 | .0025 |
| Other inputs | −.0426 | −.2909 | .2686 | .0284 | .0040 | .0146 | .0005 |

As expected, the restrictions for local separability do far less violence to the data than do those for global separability. Of the latter restrictions, the linear set are most restrictive. Again, this makes sense because they imply both output separability from inputs and input separability from outputs (i.e., input-output separability). The nonjointness hypothesis, even though it only applies locally, clearly does substantial violence to the data. This is hardly surprising since a common source of jointness in production is the use of an allocatable fixed input (Shumway, Pope, and Nash). (In our case this is land, which is a binding constraint at every sample point.) The last entry provides the system SSR for a model in which all of the second-order coefficients are set equal to zero. This is just the Cobb-Douglas case. A very large SSR confirms our earlier finding that most of these terms appear to differ substantially from zero.

### *Aggregating Outputs*

The traditional approach to econometric modeling of production for a multicrop enterprise has been to aggregate outputs. We explore the implications of this approach using a Toernqvist index, which maintains consistency between the form of the output aggregate and the translog profit function (Diewert 1976). The resulting gross elasticities are presented in table 5. A few points are immediately obvious when comparing the uncompensated elasticities in table 5 to those for the multiproduct firm in table 4. First, in place of the elastic, crop-specific supply responses, we have a relatively inelastic supply function for the crop aggregate. This is hardly surprising since land area is fixed in this data set. These results also correspond nicely to Heady's conclusions on individual crop versus aggregate supply response. A second point worthy of note is that

**Table 5. Gross Elasticities for the Output-Aggregated Model**

| | Price | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Quantity | Output | Labor | Machinery | Combine | Drying | Storage | Fertilizer | Chemicals | Other Inputs |
| Output | .3818 | −.0753 | −.0135 | −.0301 | −.0153 | −.0160 | −.0999 | −.0490 | −.0836 |
| Labor | .7563 | −.3745 | −.0297 | −.0017 | −.0137 | −.0102 | −.1344 | −.0787 | −.1065 |
| Machinery | 1.0403 | −.2277 | −.2655 | −.1065 | −.0709 | −.0203 | −.1380 | −.1002 | −.1113 |
| Combine | 1.4227 | −.0079 | −.0654 | −.5484 | −.1283 | −.1188 | −.2404 | −.1925 | −.1211 |
| Drying | 1.1728 | −.1042 | −.0705 | −.2075 | −.3060 | −.0383 | −.2430 | −.1542 | −.0503 |
| Storage | .4865 | −.0554 | −.0085 | −.0812 | −.0158 | −.1725 | −.0979 | −.0489 | −.0063 |
| Fertilizer | .6646 | −.0890 | −.0119 | −.0338 | −.0211 | −.0201 | −.3235 | −.0535 | −.1117 |
| Chemicals | .8205 | −.1311 | −.0218 | −.0681 | −.0337 | −.0253 | −.1345 | −.2665 | −.1396 |
| Other inputs | .6196 | −.0785 | −.0107 | −.0190 | −.0049 | −.0015 | −.1244 | −.0618 | −.3189 |

**Table 4. Extended**

| Price | | | |
|---|---|---|---|
| Storage | Fertilizer | Chemicals | Other Inputs |
| −.0349 | −.1625 | −.0095 | .0106 |
| .0407 | .2243 | .0400 | .0837 |
| .0462 | −.2975 | .0084 | −.4387 |
| .0077 | .0070 | .0261 | .0379 |
| .0136 | .0084 | .0315 | .0398 |
| −.0265 | .0060 | .0140 | .0853 |
| −.0326 | −.2068 | .0108 | .0049 |
| −.0268 | .0075 | .0541 | .1015 |
| .0015 | −.0806 | .0391 | .0052 |
| .0269 | .0966 | .0035 | .0491 |
| .0229 | .0058 | .0223 | −.0338 |

aggregation of outputs also affects the input demand elasticities. In this case the inputs have all become substantially more responsive to own-price changes, and all inputs are now gross complements. That is, a price increase in one input results in a decline in the demand for every other input. Apparently some of the output substitution induced by changes in relative input prices is now submerged in the estimates of input responsiveness.

Finally, recall the SSRs resulting from various output-separability restrictions. The approximate separability condition did relatively little violence to the data, whereas the global restrictions were more damaging. The elasticities in table 5 are evaluated at the point of approximation and, as such, might be expected to be somewhat plausible. However, as we move away from that point we can expect them to diverge further from those of the multiproduct profit function.

## Summary and Conclusions

In this paper we have outlined the development and use of a pseudo data set for purposes of exploring the use of a translog, multiproduct profit function. The vehicle used to generate pseudo data is a modified version of the B-9 linear programming model for a representative Indiana farm producing corn, soybeans, and wheat. It pays particular attention to the problem of timing production activities and has been extensively validated over its fifteen-year evolution.

A fractional factorial, composite experimental design was employed in creating the data set. It provides an efficient design for direct estimation of the parameters in any of the flexible functional forms corresponding to second-order Taylor approximations. Single-equation estimation of the profit function is compared to estimation via a system of share equations, and the latter is found to be preferable. Knowledge of the underlying process model permits an extensive discussion of the resulting compensated and uncompensated elasticities. This is definitely a valuable feature of using pseudo data as a teaching tool.

Several sets of restrictions on the multiproduct firm's technology are also illustrated using the pseudo data set. While classical statistical inference is not appropriate in this context, the resulting system SSRs provide a useful indication of the relative violence done by each of these restrictions. Finally, estimation of an aggregated output profit function illustrates how gross elasticities are likely to be altered when a single-product approach is employed in analyzing a problem which is inherently multiproduct in nature.

It is important to note that, by choosing to pursue a dual approach in this paper, we are unable to take advantage of information about the allocation of the fixed (land) input which is readily available in the process model's solutions. Just, Zilberman, and Hochman have suggested a primal approach which takes advantage of this information. Of course, working through the production function necessitates more restrictive functional forms, and the appropriate approach will depend on the nature of the research problem as well as the data set. As such, this may be an issue which could be usefully explored via the use of pseudo data.

In sum, there are numerous advantages to using pseudo data as a teaching tool in production economics. In particular, we feel that it enables students to focus on the particular method being taught, learning its strengths and its drawbacks. In abstracting from the inevitable problems of data quality and aggregation, it is also hoped that some of the cynicism which frequently develops with regard to the use of potentially valuable methods can be avoided.

[*Received July 1985; final revision received December 1985.*]

## References

Anderson, V., and R. A. McLean. *Design of Experiments: A Realistic Approach.* New York: Marcel Dekker, 1974.

Appelbaum, E. "Testing Neoclassical Production Theory." *J. Econometrics* 7(1978):87–102.

Berndt, E. R., and L. R. Christensen. "The Translog Function and the Substitution of Equipment, Structures, and Labor in U.S. Manufacturing, 1929–68." *J. Econometrics* 1(1973):81–114.

Binswanger, H. "A Cost Function Approach to the Measurement of Factor Demand and Elasticities of Substitution." *Amer. J. Agr. Econ.* 56(1974):377–86.

Blackorby, C., D. Primont, and R. R. Russell. *Duality, Separability, and Functional Structure: Theory and Economic Applications.* New York: American Elsevier, 1978.

Denny, M., and M. Fuss. "The Use of Approximation Analysis to Test for Separability and the Existence of Consistent Aggregates." *Amer. Econ. Rev.* 67(1977): 404–18.

Denny, M., and C. Pinto. "An Aggregate Model with Multiproduct Technologies." *Production Economics: A Dual Approach to Theory and Application,* eds., M. Fuss and D. McFadden, chap. V.1. Amsterdam: North-Holland Publishing Co., 1978.

Diewert, W. E. "Exact and Superlative Index Numbers." *J. Econometrics* 4(1976):115–45.

———. "Functional Form in the Theory of Production and Consumer Demand." Ph.D. thesis, University of California, Berkeley, 1969.

Edelman, Mark A. "Alternative Farm Commodity Program Options for the Eighties: An Analysis of Economic Implications with Emphasis on Growth of Indiana Crop Farms." Ph.D. thesis, Purdue University, 1981.

*Farm Planning and Financial Management.* Purdue University Coop. Extens. Serv. ID-68. 1980.

Griffin, James M. "Long-Run Production Modelling with Pseudo-Data: Electric Power Generation." *Bell J. Econ.* 8(1977a):112–27.

———. "The Econometrics of Joint Production: Another Approach." *Rev. Econ. and Statist.* 59(1977b):389–97.

———. "Joint Production Technology: The Case of Petrochemicals." *Econometrica* 46(1978):379–96.

———. "Pseudo-Data Estimation with Alternative Functional Forms." *Advances in Applied Microeconomics,* ed., V. K. Smith, vol. 2. Greenwich, CT: JAI Press, 1982.

Heady, E. O. "The Supply of U.S. Farm Products under Conditions of Full Employment." *Amer. Econ. Rev.* 45(1955):228–38.

Hertel, T. W. "Applications of Duality and Flexible Functional Forms: The Case of the Multiproduct Firm." Purdue University Agr. Exp. Sta. Res. Bull. No. 980, 1984.

Judge, G. G., W. E. Griffiths, R. C. Hill, and T. C. Lee. *The Theory and Practice of Econometrics.* New York: John Wiley & Sons, 1980.

Just, R. E., D. Zilberman, and E. Hochman. "Estimation of Multicrop Production Functions." *Amer. J Agr. Econ.* 65(1983):770–80.

Kmenta, J., and R. F. Gilbert. "Small-Sample Properties of Alternative Estimators of Seemingly Unrelated Regressions." *J. Amer. Statist. Assoc.* 63(1968):1180–1200.

Lau, L. J. "A Characterization of the Normalized Restricted Profit Function." *J. Econ. Theory* 12(1976): 131–63.

———. "Applications of Profit Functions"; "Testing and Imposing Monotonicity, Convexity, and Quasi-Convexity Constraints." *Production Economics: A Dual Approach to Theory and Applications,* eds., E. Fuss and D. McFadden, chap. I.3, appendix A.4, respectively. Amsterdam: North-Holland Publishing Co., 1978.

Leatham, David J., and Timothy G. Baker. "Real Capital Recovery, Changes, and Cost-Minimizing Replacement Ages of Tractors and Combines." Dep. Agr. Econ., Purdue University, March 1982.

Lopez, R. "Applications of Duality Theory to Agriculture." *West. J. Agr. Econ.* 7(1982):353–66.

———. "Estimating Substitution and Expansion Effects Using a Profit Function Framework." *Amer. J Agr. Econ.* 66(1984):358–67.

Maddala, G. S., and B. Roberts. "Alternative Functional Forms and Errors of Pseudo Data Estimation." *Rev. Econ. and Statist.* 62(1980):323–27.

McFadden, D. "Cost, Revenue, and Profit Functions." *Production Economics: A Dual Approach to Theory and Applications,* eds., E. Fuss and D. McFadden, chap I.1. Amsterdam: North-Holland Publishing Co., 1978.

McKinzie, L., T. Hertel, and P. Preckel. "Documentation of Model B-29: An Indiana Cropping Model Incorporating Endogenous Input Choices and Rotation Corn-Soybeans." Purdue University Agr. Exp. Sta. Bull. No. 461, 1984.

Pope, R. D. "To Dual or Not to Dual?" *West. J. Agr. Econ.* 7(1982):337–52.

Sakai, Y. "Substitution and Expansion Effects in Production Theory: The Case of Joint Production." *J. Econ. Theory* 9(1974):255–74.

Shephard, R. W. *Cost and Production Functions.* Princeton NJ: Princeton University Press, 1953.

Shumway, C. R. "Supply, Demand, and Technology in a Multiproduct Industry: Texas Field Crops." *Amer. J. Agr. Econ.* 65(1983):748–60.

Shumway, C. R., R. D. Pope, and E. K. Nash. "Allocatable Fixed Inputs and Jointness in Agricultural Production: Implications for Economic Modeling." *Amer. J. Agr. Econ.* 66(1984):72–78.

Statistical Engineering Laboratory. *Fractional Factorial Experiments for Factors at Two Levels.* Washington DC: National Bureau of Standards, 1957.

White, H. "Using Least Squares to Approximate Unknown Regression Functions." *Int. Econ. Rev.* 21(1980):149–70.

Young, D. L. "Relevance of Duality Theory to the Practicing Agricultural Economist: Discussion." *West. J. Agr. Econ.* 7(1982):367–72.