Nonparametric and Semiparametric Estimation: An Analysis of Multiproduct Returns to Scale

Giancarlo Moschini

The nonparametric and semiparametric regression models based on kernel density estimation are reviewed, and they are applied to the analysis of returns to scale in dairy farms using a multioutput cost function. Semiparametric tests of constant returns to scale and homogeneity of production are developed, and their implementation show that both of these hypotheses are rejected. Kernel estimation of the local degree of multiproduct returns to scale confirms that the sample's dairy farms are characterized by a considerable degree of scale economies.

Key words: cost function, dairy, flexible functional forms, kernel estimation, scale economies, semiparametric regression.

The functional specification of econometric models is a crucial component of applied research in agricultural economics that has been the object of considerable research. Notable is the introduction of flexible functional forms, pioneered by Diewert and by Christensen, Jorgenson, and Lau. This concept (see Barnett for a concise review) centers on the local approximation properties of a Taylor-series expansion of the unknown true function. Gallant and White, however, have exposed a number of statistical drawbacks of these local approximations, which has led to the development of models, such as Gallant's Fourier functional form, that can achieve global properties by the use of nonparametric approximations. This points the way to more general, and somewhat simpler, nonparametric estimation techniques.

The usefulness of nonparametric methods in applied analysis is underscored by Afriat's analysis of the restrictions of consistent choices. This approach has been extended to the analysis of production by Varian and by Chavas and Cox. In this paper, however, the term "nonparametric" applies to the statistical method of density estimation. Using this approach one can estimate the regression function or other functionals of interest without making any parametric assumptions about the true functional form. A parametric form can sometimes be retained for a portion of the model, and this leads to semiparametric specifications that can prove useful in hypothesis testing. Unlike the Afriat-type nonparametric methods, the nonparametric approach based on density estimation that is used in this paper has a solid statistical foundation, which is appealing for empirical applications requiring formal tests of hypotheses.

The nonparametric and semiparametric approaches to econometrics have attracted considerable interest in recent years (two general surveys are Ullah 1988b and Robinson 1988b). The purpose of this paper is to review the kernel method of nonparametric estimation and to apply this method to the analysis of multiproduct returns to scale using a cost function of Ontario dairy farms. Specific attention is devoted to the formulation of tests of interest in semiparametric form. The paper is organized as follows. The kernel density estimator of the regression function is reviewed, including a discussion of implementation procedures, statistical properties, and an introduction to the semiparametric regression. A multioutput cost function model is then presented. Two semiparametric tests are developed: a test of constant returns to scale, and a test of homogeneity of production. Finally, the kernel estimator of partial derivatives is used to obtain a measure of the local degree.

Giancarlo Moschini is an assistant professor of economics, Iowa State University.
Journal Paper No. J-13768 of the Iowa Agriculture and Home Economics Experiment Station Project No. 2875.
The author thanks an anonymous reviewer for helpful comments.

Copyright 1990 American Agricultural Economics Association
of multiproduct returns to scale in dairy farming.

Nonparametric Estimation

To introduce the nonparametric approach to estimation, consider a simple problem involving two continuous random variables $Y$ and $X$, and a sample of independent realizations $\{Y_1, X_1; Y_2, X_2; \ldots; Y_n, X_n\}$ from a continuous bivariate distribution function with density $f(y, x)$. The typical problem can be formulated in terms of estimating the expectation of $Y$ conditional on the explanatory variable $X$, i.e., $E(Y|X) = R(X)$, where $R(X)$ can be interpreted as the regression function that underlies the model:

$$Y = R(X) + U,$$

where $U$ is an error term satisfying $E(U|X) = 0$ by construction. Usually, $R(X)$ is assumed to have a parametric form, say the linear form $R(X) = \alpha + \beta X$. Given the sample of realizations of the variables $Y$ and $X$, one can estimate the parameters involved, thereby estimating the conditional expectation $R(X)$. Alternatively, one may note that the conditional expectation of $Y$ given $X = x$ is

$$E(Y|X = x) = \int_{-\infty}^{\infty} y \frac{f(y, x)}{g(x)} dy,$$

where $g(x) = \int f(y, x) dy$ is the marginal density of $X$. The nonparametric approach utilizes (2) to estimate $E(Y|X = x) = R(x)$ by approximating the unknown density functions $f(y, x)$ and $g(x)$ given the sample of realizations of the random variables $Y$ and $X$.

Silverman discusses the use of nonparametric methods for the estimation of density functions. Consider first the estimation of the univariate marginal density $g(x)$ using the sample of independently drawn realizations $X_1, \ldots, X_n$. From the definition of a probability density function, we have

$$g(x) = \lim_{h \to 0} h^{-1} \Pr(x - \frac{1}{2}h < X < x + \frac{1}{2}h).$$

Because a sample-based estimate of the probability $\Pr(.)$ is given by the fraction of the sample falling in the interval $(x - \frac{1}{2}h, x + \frac{1}{2}h)$, we can choose a "small" number $h$ and define the (naive) estimator $\hat{g}(x)$ of the density function as

$$\hat{g}(x) = (Nh)^{-1} \sum_{i=1}^{N} I((x - X_i)/h),$$

where $I(.)$ is an indicator function satisfying $I[w] = 1$ if $-1/2 < w < 1/2$ for $w = (x - X_i)/h$, and $I[w] = 0$ otherwise. Thus, the estimator $\hat{g}(x)$ is essentially based on the histogram constructed on the interval $x \pm \frac{1}{2}h$, where $h$ is the window width. A significant deficiency of this estimator is that it is not smooth. Also, where the data are sparse, $\hat{g}(x)$ could be assigned the value zero, which may be contrary to prior beliefs about the true density. The kernel estimator, introduced by Rosenblatt, overcomes these problems. This estimator is obtained by substituting a kernel function $k(w)$ for the indicator function in (4), that is,

$$\hat{g}(x) = \frac{1}{Nh} \sum_{i=1}^{N} k((x - X_i)/h),$$

where the kernel function $k(w)$ typically is itself a symmetric density function, i.e., satisfying $k(w) \geq 0$, $\int k(w) dw = 1$, and $\int wk(w) dw = 0$.

The kernel density estimator can be extended to the multivariate case. For instance, the joint density $f(y, x)$ is estimated by

$$\hat{f}(y, x) = \frac{1}{Nh^2} \sum_{i=1}^{N} k_2((y - Y_i)/h, (x - X_i)/h),$$

where $k_2(w', w)$ is a bivariate kernel satisfying $k(w) = \int k_2(w', w) dw'$. From (2), the kernel estimator of the regression function is obtained as

$$\hat{R}(x) = \int_{-\infty}^{\infty} \frac{\hat{f}(y, x)}{\hat{g}(x)} dy.$$

In practice it is not necessary to estimate the individual densities in (7). Using the estimators in (5) and (6) one obtains

$$\hat{R}(x) = \frac{\sum_{i=1}^{N} \int y k_2((y - Y_i)/h, (x - X_i)/h) dy}{h \sum_{i=1}^{N} k((x - X_i)/h)}.$$

or, because of the transformations $w = (x - X_i)/h$ and $w' = (y - Y_i)/h$,

Thus, the of the $Y_i$'s specific assignments that

This req

the genera

suffic

will suffi

in the d

elasticit

atives of

(12) $\hat{\beta}_j(x)$

where $(x, \ldots, X_n)$,

Implement

The implem

ors of p"
(9) \[ \hat{R}(x) = \frac{\sum_{i=1}^{N} \int (Y_i + hw') k_2(w', w) h \, dw'}{h \sum_{i=1}^{N} k((x - X_i)/h)} \]

After evaluating the integral, and noting that \( \int w' k_2(w', w) \, dw' = 0 \) for symmetric kernels, the estimated conditional mean can be rewritten as

(10) \[ \hat{R}(x) = \sum_{n=1}^{N} Y_n r_n(x), \]

where

(11) \[ r_n(x) = \frac{k((x - X_n)/h)}{\sum_{i=1}^{N} k((x - X_i)/h)} \]

Thus, the estimator \( \hat{R}(x) \) is a weighted average of the \( Y_n \)'s, where the weights \( r_n(x) \)'s are sample-specific and depend on the evaluation point \( x \), assigning the greatest weights to the observations that are closest to the evaluation point.

This regression estimator can be extended to the general case of many \( X \) variables. In the notation above, simply reinterpret \( x, X_i, \) and \( w \) as vectors of dimension \( M \), where \( M \) is the number of \( X \) variables, and substitute \( K(w) \) for \( k(w) \), where \( K(w) \) is a multivariate kernel function.

Sometimes the estimation of the conditional mean, or regression function, is the main point of interest of applied research, and the above will suffice. Often, however, one is interested in the derivatives of the regression function, which is typically the case if one focuses on elasticities. An obvious estimator for these derivatives is given by \( \hat{\beta}_i(x) = \frac{\partial \hat{R}(x)}{\partial X_i} \). Vinod and Ullah estimate \( \hat{\beta}_i(x) \) using analytic derivatives of (10), whereas Ristone and Ullah suggest the numerical derivative:

(12) \[ \hat{\beta}_i(x) = \frac{[\hat{R}(x + \frac{1}{2}h) - \hat{R}(x - \frac{1}{2}h)]/h}{\frac{1}{2}h} \]

where \((x + \frac{1}{2}h) = (x_1, \ldots, x_{j-1}, x_j + \frac{1}{2}h, x_{j+1}, \ldots, x_M)\), and \((x - \frac{1}{2}h)\) is similarly defined.

Implementing the Kernel Estimator

The implementation of the nonparametric estimators of the regression function and of the partial derivatives is essentially based on equation (10). This shows that \( \hat{R}(x) \) is a weighted average of the \( Y_n \)'s, with weights depending on the \( X_n \)'s and on the evaluation point \( x \), and as such relatively straightforward to compute. Two main choices are necessary at this stage. First, one must choose a kernel function \( K(w) \). The choice of the kernel has been shown not to be crucial (Epanechnikov). A common practice is to use the multinormal density \( K(w) = \prod_{j=1}^{M} k(w_j) \), where

(13) \[ k(w_j) = (2\pi)^{-1/2} \exp(-1/2w_j^2). \]

To accelerate the asymptotic bias reduction, Robinson (1988a) suggests the use of higher-order kernels, which can be constructed by linear combinations of (13), with the weights chosen such that the kernel has zero moments up to the desired order (see also Hong and Pagan).

The choice of the window width \( h \), on the other hand, is much more important because \( h \) controls the amount of smoothing imposed on the data. Criteria that can be used range from trial-and-error with subjective choice to a variety of automatic selection methods (Marron). The point to notice is that the variance of the kernel estimator is inversely related to \( h \), while the bias varies directly with \( h \). Thus, if \( h \) is too large, the shape of the underlying density function is obscured (causing bias); while, if \( h \) is too small, spurious detail is introduced (causing imprecision). A viable trade-off, advocated by Ullah (1988b) and Bierens, consists in choosing

(14) \[ h = c N^{-1/(4+M)} \]

where \( c \) is a proportionality constant, \( N \) is the sample size, and \( M \) is the number of nonparametric regressors involved. As in Hong and Pagan, in the empirical application that follows (14) is used and a different proportionality constant for each variable is defined, setting \( c \) equal to the sample standard deviation of the \( j \)'th \( X \) variable.2

**Statistical Properties**

The statistical properties of the kernel estimator have been analyzed extensively; the main results are reviewed in Bierens, Prakasa Rao, and Ullah (1988b). Under certain assumptions, a crucial one being that \( h \to 0 \) and \( Nh^{M} \to \infty \) as

\[ \text{This is important to account for the possible different scale of the regressors. The procedure adopted is equivalent to putting } c = 1 \text{ and linearly transforming the data to have unit covariance matrix prior to the nonparametric estimation.} \]
\( N \to \infty \), it is shown that the estimator \( \hat{\beta}(x) \) is consistent and asymptotically normal, and a similar property is established for \( \hat{\beta}(x) \). In particular, the variance of \( \hat{\beta}(x) \) can be estimated by (Uhali 1988a):

\[
V(\hat{\beta}) = \tilde{V}(x)[NH^{n+2} \hat{g}(x)]^{-1} \int (K')^2 dw,
\]

where \( \hat{g}(x) \) is given by (5), \( K' = \partial K(w)/\partial w \), and \( \tilde{V}(x) \) estimates the conditional variance of \( Y \) given \( X = x \) as

\[
\tilde{V}(x) = \sum_{s=1}^{N} Y_s^2 r_s(x) - \hat{\beta}(x)^2.
\]

It should be noted, however, that the rate of convergence of the nonparametric estimators is slower than the root-\( N \) rate of parametric models. The maximal rate of convergence in distribution of the kernel estimator is \( (NH)^{1/2} \) for the conditional mean and \( (NH^{n+1})^{1/2} \) for the response coefficients. These rates are slower than \( N^{1/2} \) because of the requirement \( h \to 0 \) as \( N \to \infty \). This feature is common to other nonparametric estimators. For example, Gallant's Fourier functional form typically does not achieve root-\( N \) consistency. When \( h \) is chosen as in (14), it is also clear that the rate of convergence of the kernel estimator is inversely related to the number of explanatory variables (the "curse of dimensionality"). For practical applications this means that large samples should be used, and that reliance on asymptotic properties for hypothesis testing may not always be appropriate.

The semiparametric model can overcome the inefficiency of the nonparametric estimators and may be more useful when the main objective is hypothesis testing.

**Semiparametric Regression**

In the semiparametric regression model of Robinson (1988a), an unknown nonlinear form is postulated only for part of the model, which can be estimated nonparametrically, while the remaining part of the model is linear (or, more generally, parametric). Let the \( Z \) denote the variables that enter the model linearly, and \( X \) denote the remaining variables. In this case the model of interest is written as

\[
Y = Z\gamma + \theta(X) + U,
\]

where \( E(U|Z, X) = 0 \). Because of the generality of \( \theta(X) \), the identification conditions for this model preclude \( Z \) from having an intercept, and require that no element of \( Z \) be included in \( X \).

In this context, interest usually centers on the estimation of \( \gamma \). If \( E(Z|X) \) is a well-defined operation, from (17) we can write

\[
Y - E(Y|X) = [Z - E(Z|X)]\gamma + U.
\]

The conditional expectations \( E(Y|X) \) and \( E(Z|X) \) in (18) can be evaluated without making unduly restrictive assumptions about \( \theta(X) \) utilizing the nonparametric regression estimator described above. Given this, \( \gamma \) can be estimated by a no-intercept ordinary least squares (OLS). Robinson (1988a) shows that the estimator of \( \gamma \) obtained in this fashion achieves root-\( N \) consistency and asymptotic normality. As will be shown below, some hypotheses typically of interest in production economics can be cast in a form such as (17), thereby providing an efficient test without making any restrictive assumption about the function being estimated.

**A Cost Function Model for Dairy Farms**

Because the nonparametric methods described above expand on the notion of flexibility that has been the tenant of many empirical applications, they offer a new approach to typical estimation and testing problems in production economics. In this paper a multiproduct cost function of Ontario dairy farms is analyzed using a large body of farm-level data. The relatively large number of observations offers a credible application for the nonparametric estimation approach, in view of the caveats on the slow rate of convergence discussed earlier. The predominantly cross-section nature of the data means that the estimated model is more suitable for studying the long-run structure of the typical farm. The analysis focuses on the issue of scale economies, which has important implications for the long-run configuration of the dairy industry in Canada.

For the problem at hand, the optimization behavior of the farms is best described conditional on the output level, which leads to a cost function representation of the production structure. More formally, if the production technology is described by the production possibilities set \( T \), which contains all the combinations \( (Q, q) \) of output vectors \( Q \) and input vectors \( q \) that are
technically feasible, given a vector \( \hat{p} \) of nominal input prices, the cost function \( C = C(Q, \hat{p}) \), defined as

\[
C(Q, \hat{p}) = \min_{q} \{ \hat{p}q : (Q, q) \in T \},
\]

provides a dual representation of this technology. Under some general conditions, \( C(Q, \hat{p}) \) is a continuous, nonnegative, nondecreasing function, positively linearly homogenous and concave in \( \hat{p} \) (McFadden).\(^4\) The property of homogeneity in prices is maintained by expressing deflated costs \( C = C/\hat{p} \), in terms of a vector of deflated prices \( p = [1, \hat{p}_2/\hat{p}_1, \ldots, \hat{p}_r/\hat{p}_1] \), where \( r \) is the number of inputs and the deflator is one of the prices (the first one in this case) such that \( C = C(Q, p) \).

This cost function is estimated using a data set consisting of 612 farm-level observations of Ontario dairy farms spread over the six years from 1978 to 1983. Inputs are aggregated into four groups: labor, feed, other intermediate inputs, and capital, and for each group a Fisher price index was constructed. The production vector consists of three output groups: milk, livestock products, and crops. Milk output is measured in hectoliters of 3.6% fat content milk, while the other two outputs represent a collection of products aggregated via a Fisher index. All right-hand-side variables are normalized to equal 1 at the median value of the sample, and the total cost of production is measured in dollars. More details on the nature of the data, and on aggregation procedures, can be found in Moschini.

Semiparametric Test of Constant Returns to Scale

The hypothesis of constant returns to scale in production represents an important benchmark. It implies that input proportions are independent of the scale of production and are determined only by the relative price of the inputs. This has important implications for the impact on input demand of supply management and other policies that may affect output at the farm level. Also, constant returns to scale justify the use of linear programming models of the farm firm, a widespread tool of farm management analysis.

Constant returns to scale are best understood in the context of homogenous production structures. Lau generalized this case to the multi-

---

\(^4\) Also, \( C(Q, \hat{p}) \) is assumed twice continuously differentiable.

---

### Table 1. Tests of Constant Returns to Scale

<table>
<thead>
<tr>
<th>Kernel</th>
<th>( \alpha )</th>
<th>Standard Error</th>
<th>( t )-Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard normal</td>
<td>0.7644</td>
<td>0.0132</td>
<td>-17.92</td>
</tr>
<tr>
<td>Higher-order</td>
<td>0.7642</td>
<td>0.0129</td>
<td>-18.21</td>
</tr>
</tbody>
</table>

\( H_0 : \alpha = 1 \)

\(^5\) Estimation was carried out on an IBM computer using user written programs coded in Fortran.
one using the normal kernel in (13) and the other using one of the higher-order kernels suggested by Robinson (1988b) (the fourth-order kernel based on (13) such that the first two moments of the kernel are both zero). In both cases, the hypothesis of constant returns to scale, \( H_0 : \alpha = 1 \), is rejected at the 5% probability level.

Because \( \alpha \) is the reciprocal of the degree of homogeneity of the production structure, it can itself be interpreted as a measure of returns to scale.\(^1\) The fact that \( \alpha < 1 \) suggests that milk production is characterized by increasing returns to scale.

**Semiparametric Test of Homogeneity**

Interpreting \( \delta \) in the previous section as the degree of multioutput scale economies may be inappropriate in that this is strictly valid only if the underlying technology is indeed homogeneous of some degree. To gain some insight into this issue, it is useful to consider other implications of the class of homogeneous production processes.

As with constant returns to scale, when the production structure is homogenous input proportions still are independent of the scale of production. This property is best expressed in terms of share equations. If \( q_j(Q, p) \) represents the input level that solves the cost-minimization problem, then \( S_j = p_j q_j(Q, p) / C \) gives the optimal cost share of the \( j \)th input, and from the derivative property \( S_j = \delta_j \ln C / \ln p_j \). If the cost function is homogenous in output, it can be written as \( C = Q^\nu C(Q, p) \). In this case \( \ln C / \ln p_j \) is independent of the output deflator \( Q \), and we can write \( S_j = S_j(Q, p) \). Thus, homogeneity in output implies that the optimal shares are independent of the scale of production.

This proposition can be tested by augmenting the share equations under homogeneity in the variable addition framework advocated by Pagan. A simple way of doing so is to write the share equations as

\[
S_j = \delta_j Q_1 + S_j(Q, p),
\]

where the linear term \( \delta_j Q_1 \) allows for nonhomo-

genous departures in a useful regression direction, that is, one in which the scale of production does influence the input shares.

The stochastic version of (23) again takes the form of the semiparametric model (17). The null hypothesis of homogeneity requires \( H_0 : \delta_j = 0 \), and under the null the model has the most flexible specification. Table 2 reports the estimated \( \delta \) coefficients for the four input cost shares. Similar to the results in table 1, no differences due to the higher-order kernels were found, and only the results obtained from the standard normal kernel (13) are reported. In all cases the estimated \( \delta \) are significantly different from zero at the 5% probability level, which leads to the conclusion that the technology of milk production is not homogenous.

The coefficients reported in table 2 also provide some insight into the scale bias in input demand. When all outputs are expanded in the same proportion, such that \( Q \) does not change but \( Q_j \) increases, the share of labor is significantly decreased. On the other hand, the share of the other three inputs increases, especially those of intermediate inputs and capital. This is consistent with the observation that cost-reducing technologies in larger-scale dairy farming typically include labor-saving solutions such as milking parlors and automatic feeding techniques. This also suggests that results on the bias of technological change from aggregate data sets ought to be interpreted carefully because scale changes are usually obscured at the aggregate level when the industry output is relatively constant and the number of firms is declining, which is the case of the Ontario dairy industry.\(^1\)

**Nonparametric Estimation of Returns to Scale**

The analysis in the previous sections showed that the cost function of Ontario dairy farms is likely

**Table 2. Tests of Homogeneity**

<table>
<thead>
<tr>
<th>Cost Share</th>
<th>( \delta_j )</th>
<th>Standard Error</th>
<th>( t )-Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labor</td>
<td>-0.0484</td>
<td>0.0039</td>
<td>-12.35</td>
</tr>
<tr>
<td>Feed</td>
<td>0.0094</td>
<td>0.0047</td>
<td>1.99</td>
</tr>
<tr>
<td>Intermediate</td>
<td>0.0273</td>
<td>0.0033</td>
<td>8.24</td>
</tr>
<tr>
<td>Inputs</td>
<td>0.0117</td>
<td>0.0044</td>
<td>2.67</td>
</tr>
</tbody>
</table>

\(^1\)This result also vindicates the contention that the alternative model need not be the most general one, as long as it explores a useful regression direction. The alternative of an homogenous structure allowed rejection of constant returns to scale, even though it is itself rejected in this section.
characterized by increasing returns to scale. Because the hypothesis of homogeneity is rejected, however, the coefficient of the semiparametric model (23) cannot be interpreted as the (constant) measure of the actual degree of scale economies. To obtain a measure of this parameter for the general cost function model, we can make use of the nonparametric estimator of partial derivatives of equation (12). Without loss of generality, the nonhomogenous cost function is rewritten with arguments in logarithmic form, i.e.,

$$\ln C = \phi(\ln Q, \ln p).$$

(24)

In keeping with the previous sections the local degree of homogeneity in output of the cost function is used as a measure of multiproduct returns to scale, that is $\phi = \Sigma_i \frac{\partial \ln C}{\partial \ln Q_i} = \Sigma_i \frac{\partial \phi}{\partial \ln Q_i}$. Thus, the degree of returns to scale involves a summation of first derivatives of $\phi(.)$, which can be estimated without any restrictive functional form assumptions by using the estimator (12).

This estimation was performed with the explanatory variables evaluated at the median point for the input prices, and at the 50%, 75%, 90%, and 95% points of the percentile distribution of output. Only the normal kernel was used, and the results are reported in table 3. Significant increasing returns to scale are found for all the evaluation points. Thus, not even the largest Ontario dairy farms seem to have exploited the existing scale economies in milk production. This conclusion is consistent with the observation that the largest farms in this industry are typically smaller than in comparable U.S. situations, and with previous findings on the U.S. dairy industry (Matulich). Note that Moschini had found locally constant returns to scale for the largest Ontario dairy farms. This difference in results is probably caused by the restrictiveness of the second-order approximation nature of the functional form used in that study (a hybrid-translog form), which underscores the desirability of the nonparametric approach.

Concluding Comments

The nonparametric kernel regression provides a way of estimating production relationships without making prior assumptions about the true shape of the function being estimated. As such, it provides an alternative to parametric flexible functional forms for econometric applications. Unlike other nonparametric methods used to study production technologies, the nonparametric and semiparametric models of this paper have a solid statistical foundation and permit hypothesis testing in a standard sense.

The application to estimating the cost function of Ontario dairy farms emphasized the derivations of simple tests of hypothesis in semi-parametric form which, under the null, do not impose undue restrictions on the shape of the underlying technology. The results indicate that constant returns to scale and homogeneity of production cannot be maintained, and that the sample's dairy farms are characterized by substantial scale economies.

[Received May 1989; final revision received October 1989.]

Table 3. Estimated Returns to Scale

<table>
<thead>
<tr>
<th>Percentile Evaluation Point</th>
<th>$\phi$</th>
<th>Standard Error</th>
<th>t-Ratio $H_0: \phi = 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>0.5489</td>
<td>0.0594</td>
<td>-7.59</td>
</tr>
<tr>
<td>75</td>
<td>0.6956</td>
<td>0.0696</td>
<td>-5.52</td>
</tr>
<tr>
<td>90</td>
<td>0.6928</td>
<td>0.1033</td>
<td>-2.97</td>
</tr>
<tr>
<td>95</td>
<td>0.7096</td>
<td>0.1352</td>
<td>-2.15</td>
</tr>
</tbody>
</table>

1. Expressing the model in terms of logarithmic variables will change the shape of the joint density of $(C, Q, P)$, but is inconsequential in our case because of the amorphous nature of the nonparametric estimator.

2. Because the measure of multiproduct returns to scale used is best interpreted as relating to the slope of ray average costs (Bau- mol, Panzar, and Willig), the evaluation point of all three outputs is the index of milk (the most important output in this case) at the given percentile levels.

References


Brown, R. S., D. W. Caves, and L. R. Christensen. "Mod-
elling the Structure of Cost and Production for Multi-


