Working Papers Series

Working Paper WP93/08 September 1993

ECONOMIES OF SIZE IN THE ONTARIO SWINE INDUSTRY

by

Alan Ker
and
Wayne H. Howard

UNIVERSITY of GUELPH

Department of Agricultural Economics and Business

University of Guelph
Guelph, Ontario
Canada
N1G 2W1

ECONOMIES OF SIZE IN THE ONTARIO SWINE INDUSTRY

by

Alan Ker
and
Wayne H. Howard

WORKING PAPER WP93/08
Department of Agricultural Economics and Business
University of Guelph

September 1993

The authors are, respectively, a graduate student, Department of Economics and Statistics, North Carolina State University (and a former graduate student, University of Guelph), and an associate professor, Department of Agricultural Economics and Business, University of Guelph.

This research was funded in part by the Ontario Ministry of Agriculture and Food and by the Ontario Pork Industry Improvement Program.

ECONOMIES OF SIZE IN THE ONTARIO HOG INDUSTRY

Compliance of sectoral policies to higher-order policies (e.g., GATT, CUSTA, and NAFTA) have recently intensified competitive pressures in Canadian and international hog/pork markets. The hog sector is particularly important in Ontario, where it accounted for approximately 20% of farm receipts between 1986-1990. Over 25% of those receipts came from exports. The ability of Ontario hog production to grow and remain competitive on international markets is in part dependent on average herd size, which grew to 241 market hogs in 1986 from 77 market hogs in 1971. Despite significant increases in herd size, the Ontario hog industry has not kept pace with farm size increases in the rest of Canada. In 1971, Ontario hog herds were 17% larger than herds in the rest of Canada. By 1986, Ontario herds were 10% smaller than those in the rest of Canada.

Economies of size and scale have become significant issues as the Ontario hog industry faces intense competition both domestically and internationally. Increasing efficiency through exploiting economies of size is seen as one way that Ontario hog producers can effectively compete on the domestic and international markets. Currently, there is a lack of information on size economies in Ontario hog production. This lack of information combined with the fact that some policies implicitly limit farm size indicates that there is a need for empirical research to determine the returns to size and scale in Ontario hog production (Gilson and Saint Louis).

The choice of method used for determining returns to size and scale has a significant influence on the level of confidence one has with one's results. Most researchers use parametric regression to estimate an average functional, such as a cost function, and then determine returns to size based upon the estimate of a local approximation to the true joint data generating process (DGP). However, the level of confidence that one can attribute to the regression estimates of the local approximation is somewhat unknown. Moreover, several studies have reported that estimation results are not robust to choice of functional form (e.g.,Chalfant; Howard and Shumway). An attractive alternative to parametric regression is nonparametric regression, which provides a consistent estimation of the joint DGP.

Nonparametric regression has its roots in electrical engineering. The procedure was developed in response to the inability to accurately estimate complicated functions using parametric regression techniques. Nonparametric regression has desirable properties that hold under relatively mild assumptions regarding the underlying distribution. Parametric estimation of functionals assumes that the parametric form is known, that the conditional variance of the dependent variable (y) given the independent variable(s) (x) is known, and usually, that the joint density distribution of y given x is normal. Unfortunately, in a parametric framework these assumptions cannot be tested, but if the assumptions are violated the results may be biased and incorrect. Moreover, the true rejection probability in hypothesis testing can be underestimated by the biased results (Ullah). In contrast, nonparametric regression only assumes smoothness and differentiability of the underlying function, enabling the data to speak for itself. The benefits of nonparametric curve estimation have been well documented (Hancoh and Rothschild, Muller, and Varian).

The objective of this study is to provide accurate quantitative information about size and scale economies in order to assist Ontario hog producers to successfully compete in the domestic and international markets. In satisfying this primary objective, three specific research objectives must be met: i) to determine if the operations from three separable production technologies (feeder pig production, farrow-to-finish, and feeder pig finishing) conform to one cost relationship; ii) to determine the level of production where size economies are fully exploited and what percentage of Ontario hog producers operate at that level; and iii) compare the results derived from non-parametric regression to those derived from parametric regression using a second-order Taylor series approximation.

This paper is organized as follows. First, the theory of returns to size and scale is discussed, including definitions and a brief review of previous studies. The parametric and nonparametric models employed in this study follows, together with a description of the data. Results from both models are presented in the final section prior to the conclusions.

RETURNS TO SIZE AND SCALE

Economies of scale is defined as the proportional change in output resulting from a proportional change in all inputs. It is commonly reported as the elasticity of scale ∅, which can be mathematically noted as

$$\varnothing = dlnf(\lambda x)/dln\lambda$$

where y = f(X) is a well-behaved (regular, monotone, and convex) production function, and λ is a scalar. To study global economies of scale requires expansion paths linear from the origin. The restriction of expansion paths linear from the origin can be relaxed to linear expansion paths to study local economies of scale (Chambers).

Economies of size is defined by the proportional change in output resulting from a proportional change in costs. Mathematically, the size elasticity is

$$\varphi = dlny/dlnc,$$

where y is output and c is total cost. The relationship between economies of size and scale is such that the elasticity of size is the envelope of the elasticity of scale in the single output case. Therefore, the two are only equivalent with respect to changes in output, if and only if a ray production technology exists (McClelland, Wetztein, and Musser). Economies of scale is related to the production technology only and does not require economic efficiency, while economies of size does require that the firm be operating on the long-run average cost curve.

Economies of size and scale consider only one output. A more realistic approach is to generalize the economies of size and scale to a multiproduct case.

Product specific size economies are defined as the proportional change in one product given a proportional change in cost, holding the other products constant. Mathematically, the incremental cost of producing product i is

$$IC_i(Y) = C(Y) - C(Y^*),$$

where $y_i \cup Y^* \equiv Y, y \notin Y^*$. Define product specific economies of size for product i, then

$$\theta_i = AIC_i/MC_i = (IC_i(Y)/y_i)/\partial C(Y)/\partial y_i.$$

If $\theta_i > 1$ ($\theta = 1, \theta < 1$) then size economies exist (size economies are fully exploited, size diseconomies exist), with respect to product i. Multiproduct size economies as defined can not be generalized to the entire product set; i.e., $IC_J = 0$, where J is the entire product set. Measuring multiproduct size economies over the entire product set requires analyzing cost with respect to a scaled unit vector of products. Mathematically,

$$\phi = \sum dlnC(Y)/dlny_i,$$

If $\phi < 1$ ($\phi = 1$, $\phi < 1$) then increasing (constant, decreasing) returns to scale exist. In effect, the measure ϕ is identifying multiproduct size economies. However, previous literature has termed this measure "multiproduct scale economies" (Baumol et al, Moschini 1990).

PREVIOUS RESEARCH

Numerous studies employing a variety of techniques and methods have been undertaken to obtain information on size and scale economies for various industries and sectors. In general, four methods are commonly cited: descriptive analysis, economic engineering, average function analysis, and frontier function analysis (Garcia and Sonka). Descriptive analysis includes comparison of average output by different firm size categories and firm survivorship as an indicator of the smallest size at which a firm may exploit economies of size. Economic engineering or the synthetic firm approach determines returns to size from different combinations of inputs/outputs. These two methods are rarely reported in the literature.

Most researchers chose either an average functional approach or a frontier function approach to model economies of size and scale. Average function analysis can be defined as fitting a function through a series of observations. The researcher can chose among production, profit, and cost functions. The production function usually has insufficient variation in the endogenous variables to obtain robust results and tends to have a relatively high level of measurement error associated with inputs such as human resources and management. The profit function is expressed in variables that are commonly exogenous and it is easier to estimate than a production function, but neither the production function nor the profit function allows measure of economies of size to be derived directly. Thus, a parametric estimation of a cost function is generally the preferred technique. There is generally sufficient variation in the endogenous variables to obtain robust results, the variables are relatively free of measurement error, and both economies of size and scale can be derived directly from the estimated cost function.

The frontier approach establishes a benchmark level of output which a firm can produce given a certain combination of inputs and technical factors. Deviations from that level of output can be due to relative inefficiency (one-sided deviations from a deterministically estimated frontier function) or due to random factors (two-sided deviations from a stochastically estimated function).

Selected studies are reported in Table 1, along with the methods employed, industry examined, and basic results. The studies reported that size and scale economies exist for the majority of industries analyzed. However, the results may be contingent on the method of analysis (e.g., Moschini, 1988; Weersink, Turvey and Godah). Hence, the choice of method for determining returns to size and scale may determine results.

METHODS

Most researchers have tended to employ either a frontier function or a parametrically estimated average functional approach to determine returns to size and scale. Recall that economies of size and scale is of interest to Ontario hog producers as part of a larger competitiveness question. The problem is not one of relative efficiency for a firm within the provincial boundaries, but one of a firm competing on an international market. The frontier approach is a relative measure that looks inward, while the average functional approach can be viewed as an absolute measure of economies of size. Therefore, the average functional approach was employed in this study. The cost function was the functional chosen for the reasons outlined above.

Additionally, one must choose between parametric and nonparametric estimation. Most researchers employ parametric regression to estimate an average functional and then determine returns to size and scale from the regression estimates. Parametric regression requires specification of a functional form that adequately approximates the joint data generating process (DGP). However, the adequacy of the approximation can not be tested given that the DGP is generally unknown. A number of local approximations (e.g., second-order Taylor, Fourier) have been employed, but results are generally not robust to functional form (e.g., Chalfant, Howard and Shumway, Vasavada and Baffes), and the adequacy of the empirically estimated approximations is unknown. Hence, the level of confidence that may be attributed to the accuracy of the results is questionable. Nonparametric regression provides consistent estimation of the joint DGP and hence yields robust results (Ullah). Therefore, nonparametric regression will be employed. However, given the numerous studies that employed parametric regression to obtain estimates of economies of size and scale, a local second-order Taylor approximation will be specified and estimated in order to compare the results from the two methods.

Parametric Model

Given a well-behaved production function Y = f(X), where Y is a 1xN vector of outputs and X a 1xM vector of inputs, a corresponding cost function exists

$$C = H(Y,W)$$

where C is cost of producing Y, and W is a 1xM vector of input prices. Expanding ln H(Y,W) in a second-order Taylor series around the means, allowing coefficients to represent derivatives, and taking a Box-Cox transformation on the set of outputs, one can obtain a Hybrid Translog function (Shumway):

$$\begin{aligned} lnC = \beta_0 &+ \sum_i \beta_i \, [(y_i^\lambda - 1) / \lambda] + \sum_j \gamma_j \, lnw_j + \\ &\frac{1}{2} * \sum_i \sum_m \alpha_{im} \, [(y_i^\lambda - 1) / \lambda] * [(y_m^\lambda - 1) / \lambda] + \frac{1}{2} * \sum_j \sum_n \delta_{jn} lnw_j lnw_n \\ &\frac{1}{2} * \sum_i \sum_j \varphi_{ij} \, [(y_i^\lambda - 1) / \lambda] * lnw_j + e \end{aligned} \quad (1)$$

Full information about the cost function can be obtained from the cost shares

$$\begin{aligned} s_j(q,w) &= \frac{\partial lnC}{\partial lnw_j} \\ &= \gamma_j + \sum \delta_{jn} lnw_n + \sum \varphi_{ij}[(q_i^\lambda - 1)/\lambda] + e_j \end{aligned} \quad (2)$$

The Hybrid Translog cost function together with the cost share equations are estimated as a system. However, given that the cost shares sum to one for every observation, the error terms sum to zero for every observation and the fully estimated system would be singular. Thus, one of the cost shares is dropped from the system and the parameters for that equation are obtained from the remaining cost shares or directly from the cost function. The estimated cost function is constrained to be consistent with the underlying theory, i.e., symmetry and homogeneity hold. Linear homogeneity in the cost function and Euler's Theorem impose the adding-up conditions of

$$\sum_j \gamma_j = 1;$$
$$\sum_j \delta_{jn} = 0; \ and$$
$$\sum_j \varphi_{ij} = 0.$$

Symmetry of the cost function requires that

$$\alpha_{im} = \alpha_{mi}, \ and$$
$$\delta_{jn} = \delta_{nj}.$$

Imposing symmetry and linear homogeneity of the cost function makes the requirement that the cost shares be homogeneous of degree zero in input prices redundant.

Nonparametric Regression

As with the parametric case, the cost model is

$$C = M(Y,W)$$

where the variables are as before, and M(Y,W) = E(C|Y,W) is an unspecified regression function. Assume there are k independent and identically distributed observations on C, Y, and W from a continuous M+N+1 variate distribution with density ϕ $(C, y_1, \ldots, y_m, w_1, \ldots, w_n)$. If $E|C| < \infty$, then

$$E(C|Y,W) = M(Y,W) = \frac{\int[C^{*}\phi(C,Y,W)]}{\phi_1(Y,W)}dC \tag{3}$$

where $\phi(C,Y,W)$ is the joint density of C given Y and W, and $\phi_1(Y,W)$ is the density of (Y,W) marginal to $\phi(C,Y,W)$. This procedure provides a value for estimated cost C. The regression coefficients necessary to obtain average and marginal incremental cost of outputs, termed response functions, are defined as the changes in C with respect to changes in a regressor, e.g., for y_j,

$$\beta(y_j) = \partial M(Y,W)/\partial Y_j. \tag{4}$$

Since $\beta(y_j)$ is a varying response coefficient, it may be valued at the mean of y_j, $[\beta(E(y_j))]$, or the expected value $E(\beta(y_j))$ may be taken to retrieve a fixed response coefficient.

Density Estimation

In order to obtain the coefficients for (4) above, densities must be estimated for (3). There are several nonparametric regression methods that could be used to estimate the densities for (3) (Muller, Silverman). Kernel estimation is generally the most intuitive of the methods and is the procedure used in this study.

Intuitive understanding of the Kernel estimator may be facilitated by looking at the histogram. The histogram is one of the oldest and simplest density estimators. Basically, a histogram tallies the number of observations that fall within an interval, termed a "bin". The bins of the histogram are defined using an origin "x_0" and bin width "h" to be $[x_0+mh, x_0+(m+1)h]$ for integers m. The intervals are closed at one end and open at the other for definiteness. Thus, the histogram estimator of (x) is

$$\hat{\theta}(x) = (nh)^{-1} * (\#\ of\ x_i\ in\ bin\ x),$$

which stratifies a set of n observations into groups or bins of width "h". Probabilities are determined from the frequency of observations belonging to a particular bin. The histogram is limited as an estimator because it is discontinuous, hence the derivatives required for equation (4) can not be obtained, and because results are subject to the choice of bin origin.

A naive estimator that eliminates the decision about bin origin can be obtained by constructing a histogram where each point is the centre of a sampling interval. By definition, if random variable x has pdf Θ , then

$$\Theta(x) = \lim_{h \to 0} (2h)^{-1} * Pr(x-h<X<x+h). \tag{5}$$

For any "h", the estimate of Pr(x-h<X<x+h) can be the proportion of the sample in the interval [x-h,x+h]. Hence, choosing a small "h", an estimate of (5) is

$$\hat{\theta}(x) = (2hn)^{-1}[\#\ of\ X\ in\ (x-h,x+h)], \tag{6}$$

A weight function can be defined such that

$$w(a) = 1\ if\ (x-h) < x_i < (x+h),$$
$$w(a) = 0\ otherwise,$$

so that (6) can be

$$\hat{\theta}(x) = n^{-1} \sum h^{-1}\ w((x-x_i)/h). \tag{7}$$

This estimator constructs a box of width 2h and height $(2nh)^{-1}$ on each observation and then sums to obtain the estimate. However, the continuity problem remains. The continuity problem in (7) can be solved by replacing w(a) with a continuous function, such as a Kernel function. A kernel function must be a continuous function satisfying

$$\int_{-\infty}^{\infty} K(x)dx = 1.$$

Replacing w(a) with K(x) in (7) yields the Kernel density estimator

$$\hat{\theta}(x) = (nh)^{-1} \sum K((x-x_i)/h). \tag{8}$$

Where (7) summed over boxes placed at each observation, the Kernel estimator places bumps at each observation and sums over the bumps. The kernel function K determines the shape of the bumps and "h", termed a smoothing parameter, determines their width. One can extend the univariate kernel to the multivariate kernel by using a d-dimensional kernel function

$$\int_{Rd} K(x)dx = 1,$$

so that (8) becomes

$$\hat{\theta}(x) = (nh^{d})^{-1} * K[(x-x_i)/h]. \tag{9}$$

When employing nonparametric kernel regression, three decisions must be made: i) choice of kernel function, ii) order of kernel, and iii) choice of smoothing parameter. For simplicity a gaussian or normal function of order two was chosen. The Epanechnikov Kernel has been shown to be the optimum function for nonparametric regression (Bendetti). However, Rosenblatt later showed that choice of a suboptimal kernel function results in only a moderate loss in efficiency as indicated by the IMSE[1]. Similarly, it has generally been found that the order of kernels do not significantly influence results (Gasser, Muller, and Mammitzsch).[2] Given that only a moderate loss in efficiency results from a suboptimal kernel function, the criteria of ease of use was used to chose the normal kernel of order two. The literature does not provide a clear choice on "h", the smoothing parameter.[3] A too small "h" yields spurious fine structure in the estimated density, while a too large "h" makes it difficult to detect the true density. A consistent, unbiased "h" was used in this study:

$$h_i = s_i/(n^{1/2})$$

$$\textit{where} \quad s^{2i} = \sum (x_i - \bar{x}_i)2/(n-1).$$

Moreover, to keep from having too much spurious detail in the tails and yet to not obscure detail in the main area of the distribution, a global smoothing parameter was used, yielding a variable kernel:

$$\hat{\theta}(x) = n^{-1} \sum (hd_{jk})^{-1} K((x-X_j)/hd_{jk}),$$

where d_{jk} is the distance from x to the k^{th} nearest data point in the set. Hence, when d_{jk} is large, the region is sparse (e.g., the tails of the distribution), and when the data points are dense, d_{jk} is small.

Nonparametric Cost Model

The nonparametric cost model is

$$C = M_n(X) + e,$$

where $M_n(X) = \Sigma C_i * r_i(X)$, and

$$r_i(X) = \frac{K_1}{\sum K_1} \frac{((x-X_i)/h)}{((x-X_i)/h)}$$

Note that e, the nonparametric residual, is robust against misspecified functional forms. For estimation of the response function a partial derivative of $M_n(X)$ is used:

Recall size economies for product Y_j are $\varphi = AIC_j/MC_j$

$$B_n(x_j) = [M_n(x_j + h/2) - M_n(x_j - h/2)]/h$$

$$= \partial M_n(X)/\ \partial x_j = C\sum (r_{1i} - r_{2i})$$

$$where \quad r_{1i} = \frac{\partial K((x-X_i)/h)/\ \partial X_i)}{\sum K((x-X_i)/h)} \tag{10}$$

$$and \quad r_{2i} = \frac{\sum (\partial k((X-x_i)/h)\ \partial x_j) * K((x-X_i)/h))}{(\sum K((x_i-x)/h))^2}$$

$$= AIC_j/\ \beta_n(y_j)$$

and scale economies are

$$\varnothing = 1 - \sum MC_{i*y_i}/C.$$

$$= 1 - \sum \hat{\beta}_{i*y_i}/C.$$

Data

Farm level data from the Ontario Farm Management Analysis Project (OFMAP) was use in this study. OFMAP is jointly administered by the Ontario Ministry of Agriculture and Food, the Farm Credit Corporation, and the Department of Agricultural Economics and Business, University of Guelph. Participation in OFMAP is generally voluntary, but some government assistance programs require participation in OFMAP. Although OFMAP has been in existence for several years, only data from 1989 and 1990 were used in this study. The data was limited to two years for maximum consistency in the collected variables, because the most recent data is thought to provide the most relevant results, and because the data set has sufficient degrees of freedom. The data set has observations on 222 farrow-finish operations, 107 feeder pig operations, and 105 finishing operations. Twenty observations from the farrow-finish sample and ten each from the feeder and finishing samples were retained for cross-validation tests. Mean values for selected sample variables are reported in Table 2.

Over 300 variables on disaggregate revenues and costs were gathered for each participating farm. Aggregation was necessary for estimation. For farrow-finish and feeder pig operations, outputs were aggregated into i) feeder pigs, ii) market pigs, iii) crops, and iv) other (including other livestock). Outputs for the finishing operations were aggregated into i) market pigs, ii) crops, and iii) other (including other livestock). The data allowed for derivation of quantities and prices for the feeder and market pigs, and crop outputs. An implicit Fisher quantity index for Other output revenues was obtained by dividing other livestock output revenues by a farm price index for livestock and livestock products (Statistics Canada), and other revenues divided by a farm price index for all farm products (Statistics Canada), and weighting the resulting indexes.

For all types of operations inputs were aggregated into labour, feed, other intermediate inputs, capital, and hog livestock. In all cases, an aggregate Fisher price index was computed and used to obtain the implicit quantity index. Labour was recorded in both total person years and hired labour expense. Hired labour expense was divided by the annual wage paid in agriculture (Ontario Ministry of Agriculture and Food) to obtain the quantity of hired labour. This quantity of hired labour was subtracted from the total person years to obtain family and owner/operator labour. An annual wage paid to family labour was inputted using the self-employed annual salary in agriculture (Statistics Canada).

The feed input aggregated crop production fed to livestock, purchased feed, and purchased forages. The tonnage of crop production fed to livestock was available from OFMAP. Prices for such crops were from Market Commentary (Agriculture Canada). Total expenses of purchased feed and forages were divided by an index of farm prices for crops (Agriculture Canada) to obtain an implicit feed quantity index.

The intermediate input aggregrated seed, fertilizer, pesticides, fuel, electricity and telephone, other crop, other animal, and other intermediate inputs. Price indices from the farm input price index (Agriculture Canada) were used to obtain an implicit intermediate inputs quantity index.

The capital input aggregated services of land, machinery, livestock herd, and buildings capital. The user cost of capital approach (Lopez) was used as these inputs are generally owned by the operator. The interest and inflation rates were those published by the Bank of Canada Review. Following Moschini (1988), depreciation was assumed to be 0.15 for machinery, 0.05 for buildings, and zero for land and livestock. To have a complete cost of capital, property taxes were added to land, and cost of repairs were added to machinery, equipment, buildings and fence capital. Animal health expenses were added to livestock capital, which was then divided by number of breeding livestock to obtain an implicit price. Similarly, land capital was divided by tillable acres to obtain an implicit land price. Input price indices for machinery and buildings (Statistics Canada) were used to obtain implicit quantity indexes.

The OFMAP data recorded the quantity of purchased hogs (market hogs, weaners and feeders, sows, gilts, and boars). Livestock prices were from the Meat and Livestock Report (Agriculture Canada).

RESULTS

Parametric Model

To determine if the three enterprise types conform to a single cost function, equations (1) and (2) were estimated as a system. Homogeneity was imposed by normalizing prices by the price of intermediate inputs. The other regularity conditions held (i.e., monotonicity and concavity). Two dummy variables (b_i, i = 1,2) were added to the system to account for the different enterprise types. It is common to test the null hypothesis that the operations from different types of enterprises conform to one cost relationship by comparing the log likelihood of the restricted (i.e., b_i=0) and unrestricted systems. However, the accuracy of the test depends on the orthogonality of the two parameters. Given that the covariance between the parameters was non-orthogonal, the joint confidence region given by

$$F[2,n-k] = = 1/2(B-b)'(S_B^2)^{-1}(B-b) \tag{11}$$

was used to test the null hypothesis. The null hypothesis is rejected at a 99% level of confidence. Hence, the three enterprise types do not conform to a common cost relationship and are modelled separately.

The estimated parameters for the three parametric models for the feeder pig, farrow-to-finish, and finishing pig enterprises are reported in Table 3. The regularity conditions hold; the average cost curves have their expect U-shape. The estimation statistics for the models are reported in Table 4. The R^2 for the cost functions are much higher than for the share equations. Low R^2's in the share equations is expected given the cross-sectional data (e.g., Moschini, 1988). Given the non-linearity of the system, the t- and f-statistics are only asymptotically valid. Diagnostic tests for heteroscedasticity and collinearity indicated the absence of those problems.

Nonparametric Model

Equations (10) were estimated with the same data as the parametric model. Nonparametric hypothesis tests are similar to parametric tests. Given that nonparametric residuals are asymptotically normal, a Likelihood Ratio (LR) can be used for hypothesis testing (Hardle):

$$LR = n*\ln(s^{2*}/s^2), \tag{12}$$

where n is number of observations, s^{2*} is the restricted nonparametric residual sum of squares, and s^2 is the unrestricted residual sum of squares. Setting the response function for each exogenous variable (analogous to parameters in parametric regression) to zero yields an asymptotically valid LR analogous to a t-test.

The overall significance of the model was determined using an Approximate Randomization Test (ART), which allows one to generate the distribution of the test statistic under the null hypothesis. The basis of the ART is that the data consists of observations (C_i,X_i), i = 1,...,n, where C is a scalar representing cost and X is a vector of input prices and output levels, and R^2 is a measure of the overall relationship between C and X. Assuming that C are iid, the null hypothesis is that the pairing of (C_i,X_i) is completely random. Simulating the data such that C_i is randomly paired with X_j, one can obtain an R^{2*} that can compared to the true R^2 from the correctly paired (C_i,X_i). This comparison produces a p-value that is analogous to the F-value in parametric regression; i.e., a high p-value suggests that the exogenous variables are not significant in predicting the endogenous variable while a low p-value suggests the opposite.

The LR was used to test whether the three types of enterprises conform to one cost function. The response functions on dummy variables representing two enterprise types were set to zero. The observed LR was 66.312; a chi-square with 2 degrees of freedom and 99% confidence is 9.21. Hence, as with the parametric estimates, the nonparametric estimates reject the null that the three enterprise types conform to a common cost function.

The estimated parameters and estimation statistics for the three models are reported in Table 5. The regularity conditions were found to hold at mean values. The high R^2 indicates a high goodness of fit. The p-values were all 0.0000, suggesting that the estimated response functions were significant in explaining cost.

Economies of Size and Scale

Economies of size and scale computed from the parametric and nonparametric models are reported in Table 6. The economies of size are reported at both "theoretical" and "practical" levels of output. Theoretical economies of size for output i occur where the average incremental cost (AIC_i) of output i equals the marginal cost of output i, or where the derivative of AIC with respect to output i is zero. If the AIC approaches its minimum relatively quickly, then the theoretical level of economies of size is a realistic recommendation for hog producers. However, if the AIC approaches its minimum over a broad range of outputs, over which increases in output correspond to infinitesimal decreases in AIC, then the theoretical level of economies of size has less practical value. "Practical Size Economies" in this study is the output level that corresponds to $dAIC/dC < 0.01$. If the AIC approaches its minimum quickly, then the "Practical" output level is in close proximity to the "theoretical" output level. However, if the AIC approaches its minimum slowly, then the "practical" level will allow the firm to exploit the great majority of the economies of size.

Elasticities of scale and size from both the parametric and the nonparametric models indicate that the great majority of Ontario swine enterprises are operating at levels that are not technical and/or economically efficient. Even when evaluated at the "practical" output level, which ranges from approximately 40-60% of the theoretically optimal level, few swine farms of any enterprise type are operating at a level that fully exploits economies of size.

Nonparametric kernel methods were used to estimate the pdf of each type of enterprise with respect to their major output (i.e., weanlings for farrowing operations, market hogs for farrow-to-finish and finishing operations). These estimated distributions are superimposed on the corresponding AIC and MC curves derived from the nonparametric models in Figures 1, 2, and 3, to provide a pictorial representation of the cost structure of the enterprises along with a the distribution of Ontario hog producers operating at the corresponding level of output. It is easy to see that the majority of hog producers of all types of enterprises could increase their efficiency as indicated by their AIC by increasing their size of operations.

COMPARISON OF PARAMETRIC AND NONPARAMETRIC METHODS

The basic results from both models are that cost functions for the three enterprise types do not conform to a single cost function and that few Ontario hog producers of any enterprise type are fully exploiting either economies of scale or size. However, the output levels where economies of size are fully exploited are different according to estimation method used. Hence, it is important to determine which of the methods most accurately models the true cost functions. Three criteria are used: estimation statistics, residual analysis, and cross-validation analysis.

Estimation statistics that indicate the overall significance of estimation for the two methods are reported in Table 7. Given that the methods produce models with different numbers of variables (e.g., the farrowing operation has 45 exogenous variables in the hybrid translog and 36 in the nonparametric model) adjusted R^2's are an appropriate comparison. For all three types of enterprises, the nonparametric models explained more variation in costs than did the parametric models. The F-values/p-values indicate that both methods reject the null hypothesis that the estimated parameters or response functions are jointly insignificant.

The residual analysis includes two tests.[4] First is an augmented non-nested Approximate Randimazation Test (ART) similar to the J-test (Davidson and MacKinnon). In this case, the ART is based on (E_i,X_i), where E_i is the residual from the parametric model and X_i is a vector of input prices and output levels. Assuming that E_i are iid, the null hypothesis is that (E_i,X_i) is random; i.e., no futher explain of the variation in costs can be explained from X. To test the null, the data is simulated such that E_i is randomly paired with X_j. The resulting (E_i,X_j) is then non-parametrically regressed, and the R^{2*} is compared to the R^2 from the correctly paired (E_i,X_i). The comparison produces a p-value analogous to the F-value. The results are reported in Table 7. The low p-valued indicate that the independent variables are jointly significant and therefore suggests that the parametric model does not adequately aproximate the actual data generating process.

The second residual analysis test is the Ansari-Bradley Distribution Free Rank test (AB) which is not dependent on degrees of freedom and does not assume normality of the residuals (Hollander and Wolfe). The AB can be used to test the goodness of fit by comparing the dispersion among the residuals from the two models. Specifically, the null is that the dispersion of the residuals from the nonparametric model is equal to or greater than the dispersal of the residuals from the parametric model. The p-values from the AB test are reported in Table 7. The low values indicate that the dispersion among the residuals from the hybrid translog models are significantly greater than the dispersion among the residuals form the nonparametric models.

A random sample of 10 observations each from the farrowing and finishing operations and 20 from the farrow-to-finish operations was selected for cross-validation. The mean absolute error (MAE), root mean squared error (RMSE), and the R^2 of the forecast for each model are reported in Table 7.[5] In all cases, the nonparametric models had lower MAE and RMSE. The magnitudes are much greater in the farrowing and finishing models than in the farrow-to-finish models. The R^2 indicate that the nonparametric model is better at out-of-sample predicting than the parametric model, though the nonparametric model does have some slippage from the in-sample R^2 to the out-of-sample R^2.

Theoretical size economies are fully exploited at greater output levels in the parametric models than in the nonparametric models, as reported in Table 6. A possible reason for this greater output level is that the estimated lambda (λ) in the hybrid translog parametric models is constrained to be equal across all outputs, and lambda represents the curvature of the cost function with respect to a particular output. Relatively small variations in output can yield an estimated cost curve that is relatively linear; such may be the case with the minor outputs (e.g., crops, other livestock). Similarly, variation around the entire range of output levels can yield a relatively non-linear cost curve; such may be the case with the major outputs (e.g., market hogs or feeders). Therefore the total incremental cost curves would be bias non-linear for the minor outputs and bias linear for the major output. The implications of a bias linear total incremental cost curve is a slow convergence to minimum AIC and hence an upward bias on the output level where economies of size are fully exploited. The convergence to minimum AIC in the Hybrid translog models was distinctly slower than in the nonparametric models. This difference in convergence rates suggests that constraining the curvature in the parametric models, i.e., _, yields relatively linear cost curves with minimum AIC at greater levels of output compared to the more "classically" shaped cost curves with minimum AIC at lower levels obtained from the nonparametric models.

SUMMARY AND CONCLUSIONS

Parametric and nonparametric estimation methods were used to analyse the current economies of size and scale in the Ontario swine industry. The parametric method employed a hybrid translog functional form. The nonparametric method employed a guassian Kernel estimator of order two. Cost functions were estimated for farrowing, farrow-to-finish, and finishing operations using 1989 and 1990 OFMAP data. Both methods rejected the hypothesis that the three enterprise types conform to a single cost function. Additionally, both methods estimated that very few Ontario hog producers are fully exploiting economies of size. In general, the results with respect to size and scale economies are quite similar. However, the parametric models do not appear to "let the data speak for itself." The nonparametric models consistently out performed the parametric models with respect to R^2, adjusted R^2, Approximate Randimazation Tests, the Ansari-Bradley test, and cross-validation results. Only the F-values/p-values were similar across models.

Perhaps the most significant finding of this paper is that the recommendations about optimum size of hog operation in Ontario is not robust to the choice of estimation method employed. The fact that the parametric method estimated minimum AIC at a greater output level than did the nonparametric method does not imply that all parametrically estimated cost curves are biased. However, there are questions about the accuracy of the recommendations about size of operations based upon parametric estimation methods. The nonparametric regression used in this study lets the data speak for itself without imposing prior restrictions about form and structure. It is possible that the restrictions imposed in order to use parametric regression have resulted in recommended size of operations that are larger than necessary to exploit economies of size.

Similar to most previous studies on economies of size, this study found that the majority of the producers in the sample could lower per unit costs of production by increasing their size of operations. Often recommendations that farmers increase their efficiency by increasing their size of operations have met with arguments from farmers that their operations are already an efficient size. Results from this study indicate that the farmers may be right. For example, the nonparametric model estimates that a farrow-to-finish operation fully exploits economies of size at 4323 market hogs per year, or 200-215 sows averaging 20-22 pigs/sow/year. However, "practical economies of size" ($dAIC/dy < 0.01$) are fully exploited in the 100-110 sows range. Given that the average farrow-to-finish operation had 83 sows in 1990 (Ontario Ministry of Agriculture and Food, 1991), and that there are several large, technically advanced hog farms with 200 or more sows, the Ontario hog industry may be more efficient than economic theory and parametric regression methods estimate.

ENDNOTES

1. The Integrated Mean Squared Error (IMSE) is a global measure of goodness of fit or accuracy. The IMSE is the integral of the Mean Square Error (MSE) which is the sum of the integrated squared bias and the integrated variance:

$$IMSE(\hat{\theta}) = \int [E\hat{\theta}(x) - \theta(x)]^2 dx + \int var\hat{\theta}(x)dx = \int E[\hat{\theta}(x) - \theta(x)]^2 dx$$

where

$$E\hat{\theta}(x) = \int (1/h) * K((x-y)/h) * \theta(y)dy,$$

$$and\ VAR\ \hat{\theta}(x) = (1/n)\int (1/h^2) * K((x-y)/h^2 * \theta(y)dy - [(1/h)\int k((x-y)/h) * \theta(y)dy]^2$$

2. A kernel is of order k if $\int K(x)*x_i dx = 0$ for all $i < k$, where i is a positive integer. Basically, higher order kernels represent higher moments in a distribution.

3. Parzen derived an optimal "h" that minimizes IMSE, but it depends on knowing the unknown density function.

4. A third test was considered but is not formally reported due to its ad hoc nature. Note that parametric regression in effect restricts the data to a weak specification by way of assuming a functional form. Nonparametric regression does not impose such a restriction. Thus, the residuals from the hybrid translog can be considered restricted and the residuals from the nonparametric model unrestricted for purposes of constructing a likelihood ratio (LR) test. Using the LR test from equation (12), with s^{2*} the restricted parametric residual sum of squares and s^2 the unrestricted nonparametric residual sum of squares, the observed LR for farrowing, farrow-to-

finish, and finishing are 100.55, 212.42, and 73.98, respectively. However, the degrees of freedom for this test have not been established. The high LR values intuitively indicate that the null hypothesis of restricting the nonparametric model to the hybrid translog model is rejected in all three cases, but one can argue that without the distribution of the test statistic that the observed LR is uninformative.

5. The R^2 for the forecast is defined as

$$R^2 = [1 - (s_a^2/s_b^2)]*100,$$

where s_a^2 is the mean squared error in predicting using the cross-validation sample and s_b^2 is the mean squared error in predicting from the estimation mean. The R^2 has a maximum of 100 but can easily be negative.

REFERENCES

Aigner, D.J. and S.F. Chu, "On Estimating the Industry Production Function:. The American Economic Review. 58(1968):826-839.

Akridge, J.T. and T. Hertel, "Multiproduct Cost Relationship for Retail Fertilizer Markets." American Journal of Agricultural Economics. 68(1986):928-938.

Ball, V.E. and R.G. Chambers, "An Economic Analysis of Technology in the Meat Products Industry." American Journal of Agricultural Economics. 64(1982):699-709.

Baumol, W.J., J.C. Panzar, and R.D. Willig. Contestable Markets and The Theory of Industry Structure. Narcourt Brace Jovanovich, New York, New York, 1982.

Bendetti, J.K. "On the Nonparametric Estimation of Regression Functions". Journal of Royal Statistical Society. Series B, Vol. 39, 1977.

Caves, D.W., L.R. Christense, and M.W. Tretheway, "Flexible Cost Functions for Multiproduct Firms". Review of Economics and Statistics. August, 1980.

Chalfant, J. A. "Comparison of Alternative Functional Forms with Applications to Agricultural Input Data." American Journal of Agricultural Economics. 66(1984):216-220.

Chambers, Robert G. Applied Production Analysis: A dual approach. Cambridge University Press, Cambridge, U.K., 1988.

Chan M.W. Luke and D.C. Mountain. "Economies of Size". Economies of Size Studies. Ames. IA: Centre of Agricultural & Rural Development, 1983.

Davidson, R. and J. MacKinnon. "Several Tests for Model Specification in the Presence of Alternative Hypotheses." Econometrica. 49(1981):406- 24.

Deller, S.C., and D.L. Chicoine, and N. Walzer. "Economies of Size and Scope in rural Volume Roads". The Review of Economics and Statistics. 70(1988):459-465.

Fare, R., L. Jansson, and C.A. Knox Lovell, "Modelling Scale Economies with Ray-Homothetic Production Functions". Review of Economics and Statistics . 67(1985):624-629.

Fleming, M.S. and I.H. Uhm, "Economies of Size in Grain Farming in Saskatchewan and the Potential Impact of Rail Rationalization Proposals". Canadian Journal of Agricultural Economics. 30(1982):1-20.

Garcia, P. and S.T. Sonka, "Methodological Issues in Assessing Economies of Size: Selected Positive Analytic Approaches". Economies of Size Studies. Ames, IA: Centre for Agricultural & Rural Development, 1983.

Gasser, T., H.G. Muller, and V. Mammitzsch, "Kernels for Nonparametric Curve Estimation". Journal of Royal Statistical Society. Series B, Vol. 47, 1985.

Hanoch, G. and M. Rothschild, "Testing the Assumptions of Production Theory: A Nonparametric Approach." Journal of Political Economy. Vol. 80, 1972.

Hardle, W. Applied Nonparametric Regression. Cambridge University Press, New York, New York. 1990.

Hock, I. "Returns to Scale in Farming: Further Evidence." American Journal of Agricultural Economics. 58(1976):745-749.

Hollander, M. and D.A. Wolfe, Nonparametric Statistical Methods. John Wiley & Sons, Toronto, Ontario, 1973.

Howard, Wayne H. and C. Richard Shumway. "Nonrobustness of Dynamic Dual Models in the U.S. Dairy Industry." Northeastern Journal of Agricultural and Resource Economics. 18(1989):18-25.

Madden, J.P. Economies of Size in Farming. AER-107. Washington, D.C.: Economic Research Service, U.S.D.A., 1967.

McClelland, J.W., M.E. Wetzstein, and W.N. Musser, "Returns to Scale and Size in Agricultural Economics." Western Journal of Agricultural Economics. 11(1986):129-33.

Moore, C.V. "Impact of Federal Limitation Policy on western Irrigated Agriculture." Western Journal of Agricultural Economics. 7(1982):301-316.

Moschini, G. "The Cost Structure of Ontario Dairy Farms: A Microeconomic Analysis." Canadian Journal of Agricultural Economics. 36(1988):187-206.

___________. "Nonparametric and Semiparametric Estimation: An Analysis of Multiproduct Returns to Scale." American Journal of Agricultural Economics. 72(1990):589-596.

Muller, H. Nonparametric Regression Analysis of Longitudinal Data. Springer-Verlag, New York, New York, 1988.

Parzen, E. "On Estimation of a Probability Density and Mode". Annals of Mathematical Statistics. Vol. 33, 1962.

Ray, S. "Translog Cost Analysis of U.S. Agricultural." American Journal of Agricultural Economics. 64(1982):490-498.

Rosenblatt, M. "Curve Estimates." Annals of Mathematical Statistics. Vol 42, 1971.

Silverman, B.W. Density Estimation for Statistics and Data Analysis. Chapman and Hall, New York, New York, 1986.

Shumway, C. Richard. 1989. "Testing Structure and Behaviour in first and second order Taylor Series Expansions". Canadian Journal of Agricultural Economics 37(1):95-110.

Ullah, A. "Nonparametric Estimation of Econometric Functionals". Canadian Journal of Economics. 21(1988):625-58.

Varian H.R. "The Nonparametric Approach to Production Analysis". Econometrica. 52(1984):579-97.

Vlastuin, C., D. Lawrence, and G. Quiggin. "Size Economies in Australian Agriculture". Review of Marketing and Agricultural Economics. 1982.

Weaver, R.D. "Discussion: On the Usefulness and Proper Measurement of Returns to Scale and Size". Economies of Size Studies. Ames I.A.: Centre of Agriculture and Rural Development. 1983.

Weersink, W., Turvey, C.G., and Godah, M. "Decomposition Measures of Technical Efficiency for Ontario Dairy Farms". Canadian Journal of Agriculture Economics. 38(1990):439-456.

Weimar, M.R., A. Hallman and L.D. Trade. "Economies of Size, Tax Reform and Profitability of Alternative Midwestern Feedlot Systems." Journal of the American Society of Farm Managers and Rural Appraisers. April, 1952.

Table 1: Survey of Selected Empirical Studies of Size and Scale Economies

Authors	Year	Approach	Method	Industry	Results
Moschini[a]	1990	Average	Nonparametric Multiproduct Cost Function	Ontario Dairy	IRS
Moschini[a]	1988	Average	Multiproduct Hybrid Translog Cost Function	Ontario Dairy	CRS
Deller, Chicoine, & Waizer	1988	Average	Multiproduct Hybrid Translog Cost Function	Rural Low-Volume Roads	IRS
Akridge & Hertel	1986	Average	Multiproduct Hybrid Translog Cost Function	Indiana & Illinois Fertilizer	IRS
Ray	1982	Average	Multiproduct Hybrid Translog Cost Function	U.S. Agriculture	IRS
Ball & Chambers	1982	Average	Non-Homothetic Translog Cost Function	U.S. Meat Products	IRS
Fleming & Uhm	1982	Average	Cost Function	Saskatchewan Grain	CRS
Caves, Christensen, & Tretheway[b]	1979	Average	Multiproduct Translog Cost Function	U.S. Railroad	IRS
Caves, Christensen, & Tretheway[b]	1979	Average	Multiproduct Hybrid Translog Cost Function	U.S. Railroad	IRS
Fare, Jansson, & Lovell	1984	Average	Homothetic Production Function	U.S. Transportation Equipment	IRS
Fare, Jansson & Lovell	1984	Average	Ray-Homogeneous Production Function	U.S. Transportation Equipment	IRS
Fare, Jansson, & Lovell	1984	Average	Ray-Homothetic Production Function	U.S. Transportation Equipment	IRS
Chan & Mountain	1983	Average	Translog Production Function	Canadian Agriculture	IRS
Vlastuin, Lawrence, & Quiggin	1982	Average	Translog Production Function	New South Wales Wheat/Sheep	IRS

Table 1: Survey of Selected Empirical Studies of Size and Scale Economies (Cont'd)

Author(s)	Year	Approach	Methodology	Industry	Results
Hoch	1976	Average	Cobb-Douglas Production Function	California Dairy	CRS
Weaver	1983	Average	Multiproduct Translog Profit Function	North & South Dakota Wheat	IRS
Yotopoulos & Lau	1971	Average	UOP Profit Function	Indian Agriculture	CRS
Weersink, Turvey & Godah	1990	Frontier	Deterministic Nonparametric Programming	Ontario Dairy	IRS & DRS
Moore	1982	Frontier	Stochastic Programming	Western U.S. Agriculture	CRS
Madden	1967	Frontier	Stochastic Programming	U.S. Beef Feedlots	IRS
Aigner & Chu	1968	Frontier	Deterministic Parametric Programming	Primary Metals	CRS
Weimar, Hallman, & Trede	1988	Synthetic	Economic Engineering	Midwest U.S. Beef Feedlot	IRS

Note: IRS = increasing returns to size or scale, CRS = constant returns, and DRS = decreasing returns.

Table 2: Mean Values for Selected Sample Variables

Variable	Enterprise Type & Year					
	Farrow-to-Finish		Farrowing		Finishing	
	1989	1990	1989	1990	1989	1990
Number Farms Reporting	110	112	57	50	59	46
Farm Size (acres)	163	161	107	121	212	208
Market Livestock Sold (Head)	751	877	78	61	825	712
Feeder Livestock Sold (Head)	144	139	1,039	1,002	0	43
Total Revenue ($)	187,469	185,605	119,833	116,826	213,101	219,808
Crop Sales ($)	16,739	11,299	15,225	13,701	44,273	27,381
Market Livestock Sales ($)	103,026	133,532	20,229	18,597	118,683	148,421
Feeder Livestock Sales ($)	21,761	10,368	47,653	59,472	4,101	6,880
Other Sales ($)	45,945	30,406	36,727	25,056	46,044	37,127
Inventory Change						
Livestock ($)	6,263	5,226	(14)	4,163	9,035	5,828
Crops ($)	3,542	4,886	1,855	4,911	6,272	5,823
Feed ($)	1,293	6,712	902	647	1,722	5,634

Table 3: Estimated Parameters for Farrowing, Farrow-to-Finish, and Finishing Operations from the Hybrid Translog Models

Coefficients	Farrowing	Farrow-to-Finish	Finishing
a	10.9360*	11.7160*	11.7070*
	(0.1028)	(0.0434)	(0.0558)
b1	0.0109	0.4956*	0.4474*
	(0.0288)	(0.0429)	(0.0471)
b2	0.4109*	0.0483	0.1035*
	(0.1024)	(0.0248)	(0.0304)
b3	-0.0654	0.0038	0.1930*
	(0.0446)	(0.0080)	(0.0470)
b4	0.3174*	0.1928	0.1191*
	(0.0917)	(0.0348)	(0.0083)
b5	0.2419*	0.1683*	0.2259*
	(0.0143)	(0.0083)	(0.0116)
b6	0.0136	0.4045*	0.3111*
	(0.0073)	(0.0129)	(0.0124)
b7	0.2431*	0.0210	0.2098*
	(0.0226)	(0.0218)	(0.0080)
b8	0.3893*	0.2451	0.0458
	(0.0210)	(0.0171)	(0.0490)
b11	-0.0135	0.0448*	
	(0.0092)	(0.0120)	
b12	0.0075	0.0326*	-0.0126
	(0.0115)	(0.0268)	(0.0157)
b13	0.0038	-0.0128	-0.0239
	(0.0036)	(0.0111)	(0.0379)
b14	0.0048	-0.1502	-0.0672
	(0.0106)	(0.0457)	(0.0062)
b15	-0.0024	-0.0480	0.0315*
	(0.0023)	(0.0078)	(0.0098)
b16	-0.0013	0.0983*	0.1445*
	(0.0006)	(0.0113)	(0.0106)
b17	-0.0044	0.0063	-0.0707
	(0.0022)	(0.0078)	(0.0063)
b18	0.0087*	-0.0416	
	(0.0025)	(0.0094)	
b22	0.0357	0.0105	
	(0.0322)	(0.0110)	
b23	-0.0035	-0.0138	
	(0.0210)	(0.0076)	
b24	-0.0122	0.0025	
	(0.0792)	(0.0253)	
b25	-0.0313	-0.0089	
	(0.0150)	(0.0045)	
b26	0.0115*	0.0151*	
	(0.0036)	(0.0065)	
b27	-0.0577	0.0030	
	(0.0132)	(0.0048)	
b28	0.0790*	-0.0047	
	(0.0129)	(0.0056)	
b33	0.0271	0.0010	0.0583*
	(0.0160)	(0.0066)	(0.0159)
b34	-0.0255	0.0149	0.0150
	(0.0240)	(0.0092)	(0.0177)

Table 3: Estimated Parameters for Farrowing, Farrow-to-Finish, and Finishing Operations from the Hybrid Translog Models (cont'd)

Coefficients	Farrowing	Farrow-to-Finish	Finish
b35	0.0104*	-0.0052	-.00055
	(0.0044)	(0.0018)	(0.0033)
b36	-0.0012	-0.0155	-0.01166
	(0.0011)	(0.0027)	(0.0052)
b37	-0.0027	0.0017	-0.0095
	(0.0035)	(0.0018)	(0.0051)
b38	-0.0201	0.0096*	0.0174*
	(0.0057)	(0.0023)	(0.0034)
b44	0.2475*	0.1929*	0.0953*
	(0.0889)	(0.0453)	(0.0456)
b45	-0.0300	-0.0162	-0.0007
	(0.0132)	(0.0071)	(0.0064)
b46	-0.0025	-0.0209	0.0333*
	(0.0034)	(0.0109)	(0.0099)
b47	-0.0222	-0.0032	-0.0566
	(0.0127)	(0.0078)	(0.0104)
b48	-0.0094	0.0204*	0.0153*
	(0.0125)	(0.0101)	(0.0064)
b55	-0.2854	0.0668	0.0439
	(0.0984)	(0.0645)	(0.0657)
b56	0.0155	0.0086	-0.0476
	(0.0213)	(0.0476)	(0.0536)
b57	0.0535	0.0529	0.0486
	(0.0739)	(0.0658)	(0.1041)
b58	0.2041*	-0.1082	-0.2792
	(0.0887)	(0.0673)	(0.0658)
b66	0.0680	-0.0802	0.0377
	(0.0368)	(0.0808)	(0.0854)
b67	-0.1488	-0.0599	-0.3179
	(0.0974)	(0.1252)	(0.1784)
b68	-0.0960	0.0485	0.1445
	(0.0704)	(0.0989)	(0.0969)
b77	-0.0023	0.2882	1.7021*
	(0.6657)	(0.6054)	(0.6313)
b78	-0.0971	0.1229	-0.7054
	(0.4322)	(0.3636)	(0.3506)
b88	-0.1104	0.2288	0.4399
	(0.3909)	(0.2979)	(0.2755)
lambda	0.0933*	0.2006	0.1494*
	(0.0540)	(0.0520)	(0.0559)

Legend: 1 - market hogs; 3- crop quantity index; 4 - other quantity index; 5 - capital price index; 6 - stock price index; 7 - labour price index; 8 - feed price index. * - asymptotically significant at $\alpha = 0.05$ level

Table 4: Estimation Statistics; Hybrid Translog Results

		Farrowing	Farrow-to-Finish	Finishing
Cost Equation	R-squared	70.31	76.52	73.45
	Adjusted R-squared	44.11	69.75	56.97
Capital Share Equation	R-squared	28.55	40.44	58.96
	Adjusted R-squared	22.05	37.97	55.14
Stock Share Equation	R-squared	24.89	41.82	26.63
	Adjusted R-squared	18.06	39.41	19.80
Labour Share Equation	R-squared	38.04	1.40	75.11
	Adjusted R-squared	32.41	-2.69	72.79
Feed Share Equation	R-squared	50.80	21.76	69.20
	Adjusted R-squared	46.33	18.52	66.33
Log-likelihood of the system		589.5233	1020.7000	530.2920
Asymptotic F - statistic of Cost Equation				
Asymptotic F - Value		2.6839	11.2977	4.4571
		0.0004	0.0000	0.0000

Table 5: Estimation Results and Statistics; Nonparametric Models

Variable	Mean(Beta)	Likelihood Ratio Test Statistic*	p-Value*
Farrowing			
Market Hogs	44.8538	14.5670	0.0001
Feeders	30.1787	79.7839	0.0000
Crop Quantity Index	5493.5317	18.9661	0.0000
Other Quantity Index	33175	99.2760	0.0000
Capital Price Index	33195	0.5998	0.4387
Intermediate Inputs Price Index	−239930	1.8678	0.1717
Stock Price Index	249890	2.5369	0.1112
Labor Price Index	−133450	0.8340	0.3611
Feed Price Index	−54286	0.1815	0.6700
Farrow-to-Finish			
Market Hogs	68.5982	243.6106	0.0000
Feeders	24.4842	49.5816	0.0000
Crop Quantity Index	5233.1665	5.9134	0.0150
Other Quantity Index	35574	184.0723	0.0000
Capital Price Index	−13831	1.5769	0.2092
Intermediate Inputs Price Index	212550	7.9488	0.0048
Stock Price Index	−24958	1.6609	0.1975
Labor Price Index	418820	7.7248	0.0054
Feed Price Index	−196920	0.8890	0.3457
Finish			
Market Hogs	91.0309	88.7610	0.0000
Crop Quantity Index	21919	54.0288	0.0000
Other Quantity Index	37836	49.3895	0.0000
Capital Price Index	2078	20.5818	0.0000
Intermediate Inputs Price Index	513280	3.2787	0.0704
Stock Price Index	−40211	1.3065	0.2530
Labor Price Index	−38642	0.3055	0.5804
Feed Price Index	−625510	2.2066	0.1374

Legend: * – asymptotically valid only

Estimation Statistics

Farrowing

R-Squared		90.39
Adjusted R-Squared		89.40
Standard Randomization Test (p-value)	0.000	

Farrow-to-Finish

R-Squared		91.73
Adjusted R-Squared		91.34
Standard Randomization Test (p-value)	0.000	

Finishing

R-Squared		84.10
Adjusted R-Squared		82.62
Standard Randomization Test (p-value)	0.000	

Table 6: Scale and Size Economies from the Estimated Paramtres and Nonparameters Models.

	Farrowing		Farrow-to-Finish		Finishing	
	Hybrid Translog	Non-parametric	Hybrid Translog	Non-parametric	Hybrid Translog	Non-parametric
Elasticity of Scale						
- at means	0.7993	0.8027	0.8510	0.7555	0.9581	0.7461
- output where scale elasticity = 1.0	2598	1805	1207	1453	1154	1327
- percent of population realizing economies of scale	6.48%	20.55%	30.39%	20.55%	25.96%	21.36%
Elasticity of Size						
- at means	1.6689	1.8227	1.4120	1.6799	1.4117	1.7873
Theoretical						
- output where size elasticity = 1.0	**	2911	4657	4323	5250	2820
- percent of population fully exploiting size economies	0.00%	3.32%	0.00%	0.00%	0.00%	1.60%
Practical						
- output where size elasticity = 1.0	1997	1543	1810	2191	2026	1790
- percent of population fully exploiting size economies	16.16%	28.83%	12.34%	6.93%	8.82%	12.35%

Legend: ** - output level exceeds an adhoc boundary which it is assumed the estimated relationship no longer holds.
- Theoretical size ecoomies refer to where the derivative of the average incremental cost curve equals 0.00.
- Practical size economies refer to where the derivative of the average incremental cost curve is less than 0.01.

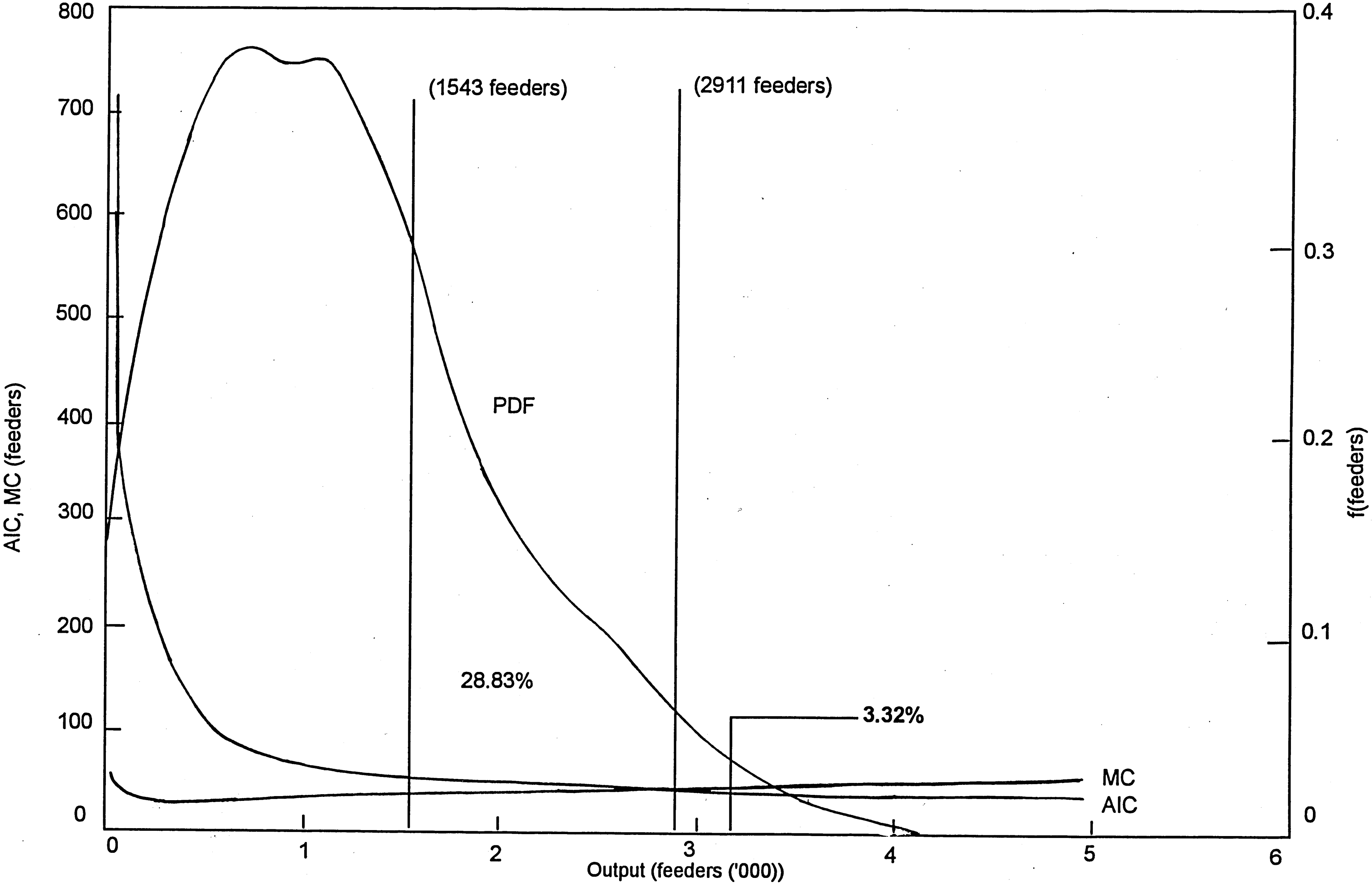
Table 7: Summary of Estimation Statistics

Statistic	Farrowing		Farrow-to-Finish		Finishing	
	Hybrid Translog Model	Nonparametric Model	Hybrid Translog Model	Nonparametric Model	Hybrid Translog Model	Nonparametric Model
R-Squared	70.31	90.39	76.52	91.73	73.45	84.10
Adjusted R-Squared	44.11	89.40	69.75	91.34	56.97	82.62
F-Value	0.0004		0.0000		0.0000	
p-Value		0.0000		0.0000		0.0000
ART p-value	0.0031		0.0058		0.0031	
Ansari-Bradley Test	0.0000		0.0000		0.0090	
Mean Absolute Error	58966	26881	30912	29548	286823	49309
Root Mean Squared Error	67733	32138	42147	36262	639122	80417
R-squared of Forecast	-134.48	47.21	64.58	73.78	-997.42	82.63

Note: R-squared = 1-(MSEx/MSEc) where MSEx is the MSE of forecasted values when the cross validation sample is put into the regression function and MSEc is the MSE of forecasted values when the mean cost from the estimation sample is the forecasted value

Figure 1

Average Incremental Cost, Marginal Cost, and Probability Distribution of Feeders
Feeder Pig Production: Nonparametric Results

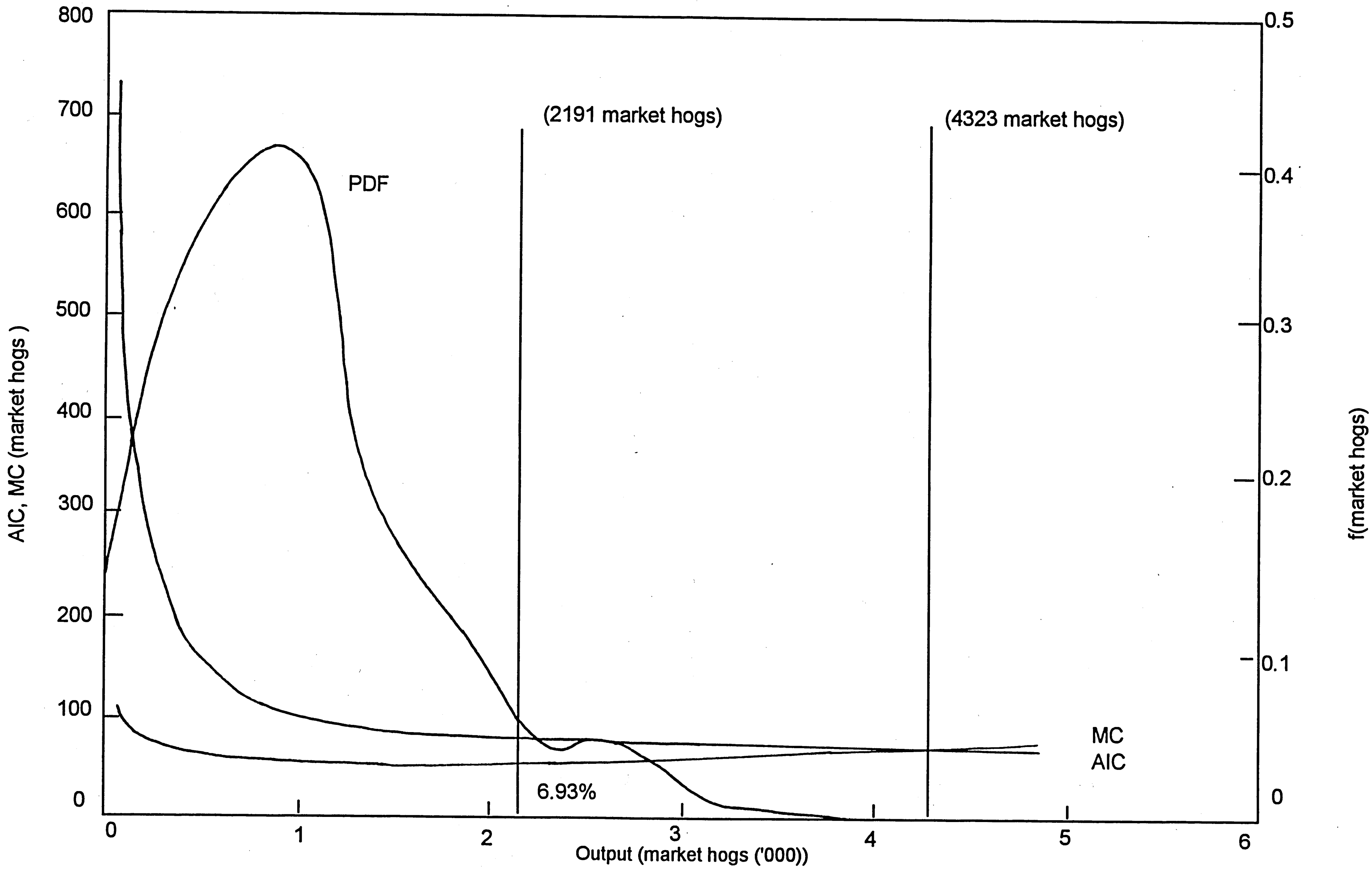


Legend: PDF - probability distribution function of feeders in feeder pig production operations in Ontario.
: MC - the marginal cost of producing a feeder in a feeder pig production operation.
: AIC - the average incremental cost of producing a feeder in a feeder pig production operation.

Note: Marginal Cost and Average Incremental Cost are evaluated at mean prices.

Figure 2

Average Incremental Cost, Marginal Cost, and Probability Distribution of Market Hogs
Farrow-to-Finish; Nonparametric Regression

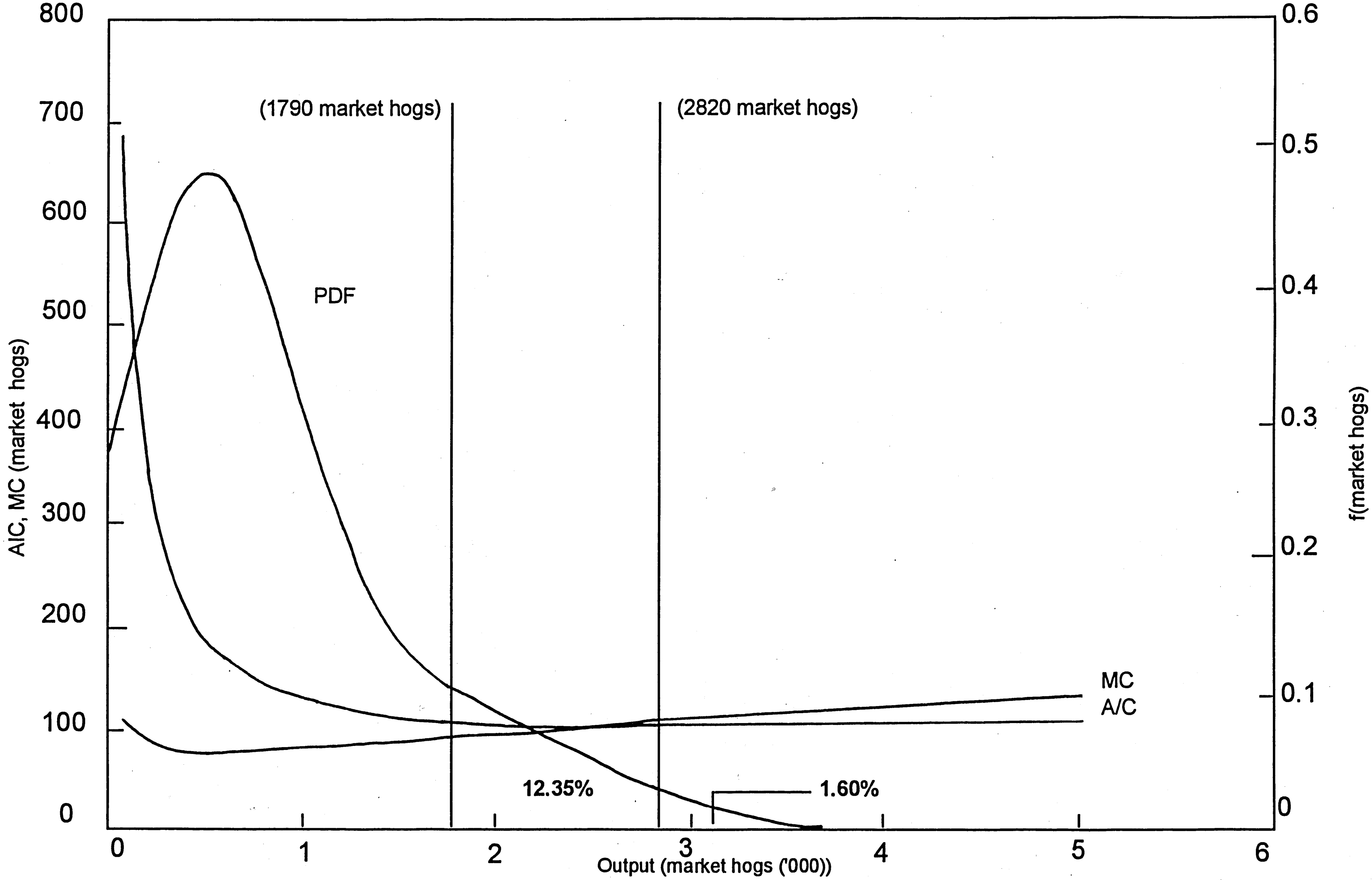


Legend: PDF- probability distribution function of market hogs for farrow-to-finish operations in Ontario.
: MC - the marginal cost of producing a market hog in a farrow-to-finish operation.
: AIC - the average incremental cost of producing a market hog in a farrow-to-finish operation.

Note: Marginal Cost and Average Incremental Cost are evaluated at mean prices.

Figure 3

Average Incremental Cost, Marginal Cost, and Probability Distribution of Market Hogs
Feeder Pig Finishing: Nonparametric Results



Legend: PDF - probability distribution function of market hogs for feeder pig finishing operations in Ontario.
: MC - the marginal cost of producing a market hog in a feeder pig finishing operation.
: AIC - the average incremental ocst of producing a market hog in a feeder pig finishing operation.

Note: Marginal Cost and Average Incremental Cost are evaluated at mean prices.