The 83rd Annual Conference of the Agricultural Economics Society

Dublin

30th March to 1st April 2009

A Nonparametric Kernel Representation of the Agricultural Production Function: Implications for Economic Measures of Technology

By Grigorios T. Livanis, Matthew J. Salois*, and Charles B. Moss

March 10, 2009†

Abstract

The issue of production function estimation has received recent attention, particularly in agricultural economics with the advent of precision farming. Yet, the evidence to date is far from unanimous on the proper form of the production function. This paper reexamines the use of the primal production function framework using nonparametric regression techniques. Specifically, the paper demonstrates how a nonparametric regression based on a kernel density estimator can be used to estimate a production function using data on corn production from Illinois and Indiana. Nonparametric results are compared to common parametric specifications using the Nadaraya-Watson kernel regression estimator. The parametric and nonparametric forms are also compared in terms of describing the true technology of the firm by obtaining measures of the elasticity of scale and the marginal physical product through nonparametric estimation of the gradient of the production surface. Finally, the elasticities of substitution are compared between both parametric and nonparametric representations.

Keywords: nonparametric regression, nonparametric derivatives, Gaussian kernel, optimization techniques, production function.

JEL Classification: C14, C15, C16, C61, Q12

*Corresponding author (m.j.salois@reading.ac.uk). Grigorios T. Livanis is Lecturer, Information, Operations and Analysis, at Northeastern University, USA. Matthew J. Salois is Research Fellow, Department of Agricultural and Food Economics at the University of Reading, UK. Charles B. Moss is Professor, Food and Resource Economic Department, at the University of Florida, USA.

†Presented at the 83rd Annual Conference of the Agricultural Economics Society, Dublin 30 March - 1 April, 2009.

1 Introduction

Contributions in duality theory have alleviated the need to specify primal production functions in the economic literature, allowing researchers to rely on optimizing behavior to estimate factor demand and output supply functions. However, the issue of production function estimation has received recent attention, particularly in agricultural economics. Revived interest in the primal production function stems not only from a more recent understanding that output supply and input demand is most efficiently derived from the empirical production function (Mundlak 1996), but also stems from the needs of the agricultural economist. New agricultural technologies in precision farming empowers farm managers with micro-level field management, permitting inputs to be applied at varying levels to very small plots of land. Such decisions can only be guided by specification of the primal (Moss and Schmitz, 2006).

While polynomial forms such as the quadratic, translog, and square-root have been used since Heady and Dillon's seminal work, *Agricultural Production Functions* (1961), an effort to combine agronomic and ecologic principles to crop yield response and production has led to interest in von Liebig's "law of the minimum" and the aptly named von Liebig production function(.[1] Yet, the evidence to date is far from unanimous on the proper form of the crop response function, and the literature still remains unsettled on the best practice for estimating agricultural production functions, which has implications for farm profitability and environmental management.

For example, the primary economic benefit of precision agriculture is more efficient use of inputs resulting in lower costs and higher profits. However. the profitability of precision agriculture is contingent upon the ability to render accurate, timely, and quality information to the decision maker. One aspect of this information is in the form of crop response models where yield is determined from an equation relating yield response to fertilizers, typically estimated by the primal. Polynomial functional forms of crop response model are often to blame for recommending inefficient use of inputs (Ackello-Ogutu et al., 1985).[2]

There are also environmental implications of inefficient fertilizer use that result from incorrectly specified crop response models. Reduced application of

[1] The von Liebig hypothesis combines the dual notions of non-substitutability of inputs and yield plateaus into production and crop response analysis and has been investigated rigorously by Quirino Paris and colleagues, as well as others.

[2] There is evidence that farmers over-fertilize their crops, for example, over application of nitrogen to corn crops seems to be common (Babcock, 1992).

nitrogen, for example, will result in decreased environmental run-off and nitrogen pollution, which are harmful to local ecosystems. These benefits can only be realized through correct specification of the primal production function. Any mistakes made in economic theory to derive farm management recommendations, such as incorrect fertilizer applications based on a false functional form, can be costly to the farmer and to the environment.

This paper addresses the problem of specifying a functional form for the production function by using nonparametric kernel regression. The kernel approach avoids the reliance on any parametric form and offers a simple method of estimation.[3] Specifically, a nonparametric kernel regression based on a multivariate Gaussian kernel is used to estimate a production function for corn in Illinois and Indiana. Nonparametric results are compared to common parametric specifications using the Nadaraya-Watson kernel regression estimator. The parametric and nonparametric forms are also compared in terms of describing the true technology of the firm by obtaining measures of the elasticity of scale and the marginal physical product through nonparametric estimation of the gradient of the production surface. Finally, the elasticities of substitution are compared between both parametric and nonparametric representations.

Section 2 section briefly reviews the literature on crop response models and the continuing debate regarding the appropriate functional form. Recent uses of nonparametric methods to similar economic problems are also discussed. Section 3 develops the nonparametric kernel regression estimator and emphasizes nonparametric estimation of the derivatives, which is required for comparing elasticities. Section 4 applies the nonparametric estimator, as well as common parametric approaches, to data on corn production and obtains estimates of the derivatives. Estimated results are compared between parametic and nonparametric methods. The final section concludes and summarizes the result.

2 Crop Response Debate

There has been a concerted effort in the agricultural economics literature in trying to obtain an appropriate functional form for production and crop response estimation. The work by Heady and Dillon (1961) promulgated the use of polynomial forms for agriculture production models for several decades. It was not until Lanzer and Paris (1981) that a serious account of ecologic and

[3]The kernel approach is well documented in the economic literature. Hardle (1990) is a classic reference. Kernel methods have been developed further by Pagan and Ullah (1999) and most recently by Li and Racine (2007).

agronomic principles were introduced into the production estimation problem through the von Liebig hypothesis. Using a linear response and plateau (LRP) function to model the "law of the minimum," the authors found that yearly fertilizer application rates could be reduced by 20 percent for the wheat-soybean double cropping system in Southern Brazil.

Using non-nested hypothesis tests, Ackello-Ogutu, Paris, and Williams (1985) rejected the square root and quadratic forms in favor of the von Liebig. The non-substitution hypothesis was later relaxed in Frank, Beattie, and Embleton (1990) which tested the Mitscherlich-Baule (MB) form against the von Liebig and quadratic. The MB allows for both factor substitution and plateau growth. Overall, the authors recommend the use of the MB form based on pairwise J-tests and P-tests.

Berck and Helfand (1990) show that differentiable production functions are not mutually exclusive from von Liebig forms and reconcile the two via dynamic calculus. The authors conclude that the quadratic form is still adequate for yield prediction goals, however, neither forms appear to better apt than the other to estimate yield changes resulting from input level changes. In response to Frank et al. (1990), Paris (1992) estimates a non-linear von Liebig model against an MB, quadratic, square-root, and linear von Liebig. Non-nested hypothesis tests seem to favor the nonlinear von Liebig over other specifications, achieving the highest level of revenue with the smallest amount of fertilizer.

Chambers and Lichtenberg (1996) point out the relatively low power of non-nested hypothesis tests and utilize a nonparametric, mathematical programming technique to determine the existence of yield plateaus and input substitutability. This study was one of the earliest attempts to incorporate nonparametric techniques to the crop response model debate. Using a dual cost function approach, Chambers and Lichtenberg find evidence of yield plateaus but also find existence of input substitutability, thus concluding the von Liebig-Paris approach is only appropriate under certain circumstances and for particular crops.

Berck, Geoghegan, and Stohs (2000) use a nonparametric data-envelope-analysis method to test the validity of the von Liebig production function and find it to be a poor fit, citing little evidence for right-angle isoquants. Holloway and Paris (2002) combine frontier methods with the von Liebig methodology using Bayesian techniques. The authors acknowledge recent nonparametric tests rejecting the von Liebig model, but are unable to reconcile results from parametric and nonparametric methods leaving the "endeavor for future research," (Holloway and Paris, 2002). More recently, Moss and Schmitz (2006) develop a semi-parametric estimator using a univariate form of the Zellner production function and represents the only study to date to incorporate nonparametric kernel techniques to crop response analysis.

The use of kernel estimation methods is not new to the applied economics, but has only recently been used to tackle the functional form problem. Moschini (1990) uses a semi-parametric normal kernel to estimate U.S. meat demand. Ker and Goodwin (2000) compare a normal nonparametric kernel estimator to the Bayesian kernel to obtain accurate crop insurance rates. Cooper (2000) compares nonparametric, semiparametric, and parametric estimators of recreational demand analysis and finds the choice of form largely depends on sample size. Daniel Henderson and colleagues have recently applied nonparametric kernel techniques to a host of economic issues, including random effects estimation (Henderson and Ullah, 2005), public and private capital productivity (Henderson and Kumbhakar, 2006), hedonic price function estimation (Henderson, Kumbhakar, and Parmeter, forthcoming), pollution abatement and foreign direct investment (Henderson and Millimet, 2007), and child health (Henderson, Millimet and Parmeter, forthcoming).

3 Nonparametric Kernel Regression

The empirical model is a multivariate regression of a $n \times 1$ scalar dependent variable y on a $n \times q$ dimensional independent variable matrix x given by:

$$E[y|x] = m(x) + \varepsilon = m(x_1, ..., x_q) + \varepsilon \tag{1}$$

where $i = 1, ..., n$ indicates the number of observations, q indicates the number of independent variables, and ε denotes an error term. A parametric regression of Equation 1 would construct a relationship between the variables by specifying a functional form for $m(x)$, based on some set of parameters, say $m(x, \beta)$. This relationship could be either a linear or nonlinear function of a finite number of parameters, given by the q-dimensional vector β.

One common functional form include the Cobb-Douglas, $y = \alpha x_1^{\beta_1} \cdots x_q^{\beta_q}$, which is given in logarithmic regression form by:

$$\ln y = \alpha + \beta_1 \ln x_1 + ... + \beta_q \ln x_q + \varepsilon \tag{2}$$

Another common functional forms is the Translog:

$$\ln y = \alpha + \sum_{i=1}^{n} \beta_i \ln x_i + 0.5 \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} x_i x_j + \varepsilon \tag{3}$$

Once functional forms such as the Cobb-Douglas in Equation 2 and Translog in Equation 3 are estimated, parameter estimates are used to test various economic measures. For example, the Cobb-Douglas production function exhibits constant returns to scale if $\sum \beta_i = 1$.

Nonparametric regression, on the other hand, constructs a relationship between y and x based on weighted averages, and thus avoids having to specify a specific functional form. These weights are inversely proportional to the distance between a given sample point, x_i, and an approximation point, x. Following Pagan and Ullah (1999), a general class of nonparametric regression estimators of Equation 1 can be written as:

$$\hat{m}(x) = \sum_{i=1}^{n} w(x_i, x) y_i \tag{4}$$

where $w(x_i, x)$ represents the weight assigned to the i^{th} observation y_i, with the weight depending on the distance of the sample point, x_i, from the point, x. The point of approximation, x, is often taken to be a point in the sample, say x_p, in which case the summations should delete the p^{th} observation. Such estimators are often referred to as "leave one out" estimators.

A number of nonparametric estimators can be defined from Equation 4, depending on how the weighting function, $w(x_i, x)$, is specified. One particular class of estimators is the Nadaraya-Watson (NW) kernel regression estimator, which defined the weighting function as:

$$w(x_i, x) = \frac{K\left(\frac{x_i - x}{h}\right)}{\sum_{i=1}^{n} K\left(\frac{x_i - x}{h}\right)} \tag{5}$$

where $K(\cdot)$ is a multivariate kernel and h is a smoothness parameter called the bandwidth that determines how many x_i's around x are used in the kernel function. Defining $\varphi_i = \left(\frac{x_i - x}{h}\right)$ and plugging Equation 5 into Equation 4 the NW kernel regression estimator is obtained:

$$\hat{m}(x) = \frac{\sum_{i=1}^{n} \left[K(\psi_i) y_i\right]}{\sum_{i=1}^{n} K(\varphi_i)} \tag{6}$$

The NW estimator takes the average of the observed y_i values and then weights the average by the chosen kernel $K(\cdot)$, which is then normalized by the sum of the weighted averages. There are various forms of the kernel to choose from. However, choice of kernel is not a critical decision, with different kernels yielding similar results (Pagan and Ullah 1999). Rather, the kernel choice often comes down to criteria of differentiability, continuity, and other characteristics. This study employs the multivariate standard normal Gaussian kernel:

$$K(x, x_i, h) = \left(\frac{1}{h\sqrt{2\pi}}\right)^{-k} \exp\left[-\frac{1}{2}(x_i - x)\, h^{-\frac{1}{2}}\, (x_i - x)'\right] \tag{7}$$

Several alternative forms of the kernel function are available, such as the uniform, Epanechnikov, and product Gaussian kernel.

A far more important decision than the kernel is the choice of bandwidth. The bandwidth parameter, h, controls the rate at which the weight given to points decline as x_i departs from the point of approximation. In the limit as $\delta \to 0$ the nonparametric regression simply becomes the average at each point. As $\delta \to \infty$, the nonparametric regression simply becomes a constant. While plug-in methods have been devised and are often used, data driven methods, such cross-validation techniques, remain the preferred way of selecting bandwidths. In the multivariate case, a different bandwidth is chosen for each independent variable, or alternatively, the data can be normalized by the standard deviation of each variable, in which case the same bandwidth can be used for each variable and h is a scalar[4].

The most common data-driven way to select bandwidths is by least-squares cross-validation (LSCV), which involves estimating the optimal bandwidth $h^\star$ from the following optimization problem:

$$h^\star = \min_h \left[\sum_{i=1}^{n} (y_i - \hat{m}_{-1}(x_i))\right]^2 \tag{8}$$

where $\hat{m}_{-1}(x_i)$ is the "leave one out" NW estimator of $m(x_i)$ obtained from the kernel formula in Equation 6. The obtained $h^\star$ from Equation 8 minimizes the asymptotic integrated mean squared error (AIMSE) . Other method of obtain the optimal bandwidth exist, such as likelihood cross-validation and the more recently developed AIC_c approach based on the Akaike information criterion. However, LSCV remains a relatively simple and common method and so is the approach taken in this paper to obtain the optimal bandwidth.[5]

[4]This paper utilizes the single scalar approach by standardizing the inputs by their standard deviation, which simplifies the estimation

[5]The interested reader is referred to Ullah and Pagan (1999) and Li and Racine (2007)

Often, interest is focused on the marginal effect of an independent variable, $x_1, ..., x_q$ has on the dependent variable y. In parametric regressions, this is usually represented by the estimated parameters, $\beta_1, ..., \beta_q$. In nonparametric regression, coefficients are actually estimated, but rather the marginal effects are obtained from nonparametric estimation of the gradient, and are often referred to as response coefficients. Pointwise derivatives can be obtained for any given value of x using either analytical or numerical methods, with such derivatives describing the local behavior or shape of the regression function.[6]

Numerical approximation of the first and second derivatives has been investigated by Ullah (1988a, 1988b), Rilstone and Ullah (1989), and Rilstone (1990) using perturbation by finite differences. While equivalent to the analytical approaches proposed by Gasser and Muller (1984) and used by others, numerical methods are easier to implement are far more computationally tractable, especially for higher order derivatives.. Following the Rilstone and Ullah, the estimated gradient is obtained from:

$$\hat{\nabla}_j^{(1)}(x) = \frac{\hat{m}(x + e_j h) - \hat{m}(x)}{h} \tag{9}$$

where e_j is a $q \times 1$ vector with unity in the j^{th} position and zeros elsewhere, which implies that $\hat{m}(x + e_j h) = m(x_1, ..., x_j + h, ..., x_q)$. Equation 9 can be equivalently written in terms of the individual response coefficients as:

$$b_j(x) = \frac{\hat{m}\left(x + \frac{1}{2}e_j h\right) - \hat{m}\left(x - \frac{1}{2}e_j h\right)}{h} \tag{10}$$

The estimated Hessian is obtained similarly from:

$$\hat{\nabla}_{jk}^{(2)}(x) = \frac{\hat{\nabla}^j(x + e_k h) - \hat{\nabla}^j(x)}{h} \tag{11}$$

which is a $q \times q$ matrix of estimated second-order and cross-partial derivatives. Using Equation 11 the partial derivatives can be seen to be equal to:

for a more detailed discussion of LSCV as well as other methods of obtaining the optimal bandwidth.

[6]Another benefit of nonparametric regression is the ability to obtain average derivatives, which describe the global curvature of the function. Although average derivatives lack the intuitive interpretation of pointwise derivatives, they offer greater statistical consistency and converge faster than pointwise derivatives. This paper focuses on pointwise derivative since clear interpretation is better suited for elasticity computation. The interested reader is referred to Pagan and Ullah (1999) and Rilstone (1991) for more information on average derivatives.

$$b_{jk}(x) = \frac{\hat{m}(x + (e_j + e_k)h) - \hat{m}(x + e_j h) - \hat{m}(x + e_k h) + \hat{m}(x)}{h^2} \tag{12}$$

The estimation of the derivatives follows from the actual estimation of the regression function, and faces the same choices regarding kernel form and bandwidth selection. Likewise, the choice of kernel is not a largely important matter, as long as higher order derivatives of the kernel exist. Kernels with higher order derivatives help to reduce any finite sample bias in the estimated derivative. For this reason, researchers who use a kernel in regression estimation that has a low order of derivatives, such as the uniform, would select a different kernel for the estimation of the derivatives.

The bandwidth selection is again the more salient factor, which must be chosen to minimize the mean square error. Following Pagan and Ullah (1999), the optimal bandwidth for pointwise derivative estimation is:

$$h^\star \propto n^{-\frac{1}{(q+4+2s)}} \tag{13}$$

which is proportional to the optimal bandwidth for the kernel regression found using the LSCV procedure in Equation 8 with q independent variables and a s^{th} order of derivative.

The finite sample and asymptotic properties of the derivative estimators in Equation 9 and Equation 11 have been investigated by Ullah (1988a, 1988b), Rilstone (1990), and Pagan and Ullah (1999). Namely, the estimators are weakly consistent and asymptotically Normally distributed as

$$\left(nh^{q+2s}\right)^{1/2}\left(\nabla^{(s)} - \nabla^{(s)}\right) \sim N\left(0, \zeta(x)\right) \tag{14}$$

where s indicates the order of the derivative and the covariance matrix $\zeta(x)$ is consistently estimated from:

$$\frac{\hat{\sigma}^2}{\hat{f}(x)} \int_{\Re^q} \left[\left(\frac{\partial^{(s)} K(\psi)}{\partial \psi^{(s)}}\right)^2 d\psi\right] \tag{15}$$

with the estimated variance given by:

$$\hat{\sigma}^2 = \frac{1}{n}\left(y - \hat{m}(x)\right)'\left(y - \hat{m}(x)\right) \tag{16}$$

and the estimated marginal density given by:

$$\hat{f}(x) = \frac{1}{nh^q} \sum_{i=1}^{n} K\left(\frac{x_i - x}{h}\right) \tag{17}$$

Hence, the joint distribution of the response coefficients are found to be normal with zero mean vector and variance matrix, $\zeta(x)$, which can be consistently estimated using Equation 15 through Equation 17, from which the standard errors of the derivative estimates are obtained.[7]

A key benefit to nonparametric estimation of derivatives is that the nonparametric derivative estimators are a function of x. Thus, a fixed response coefficient could be defined at a particular point of x, such as the mean, to give an estimate of $b_j(\bar{x})$, or could likewise be evaluated at the median. Alternatively, one could evaluate the response coefficients over every value of x, which gives a great advantage over parametric estimation. Namely, a local inflection point of the function $m(x)$ could occur anywhere in the sample and so evaluating the derivative at the mean ignores the possibility that the function may not be well-behaved at that point or in a neighborhood around that point.

The estimates of the response coefficients in the gradient gradient and the second-derivatives in the Hessian can then be used to compute common measures of economic efficiency and production technology, such as output elasticities, marginal product of inputs, and elasticities of scale and substitution. The output elasticities measures the percent change in output given a one percent change in a given input and is given by:

$$\epsilon_i = \frac{\partial \ln m(x)}{\partial x_i} = \frac{\partial m(x)/\partial x_i}{x_i/m(x_i)} \tag{18}$$

where the elasticity of scale is simply the sum of the output elasticties ($\epsilon_i = \sum_i \epsilon_i$) and measures the percent change in output given a one percent change in all inputs. The elasticity of scale and the output elasticties tells us there is a relationship between returns to scale and marginal productivity. Since the marginal product is $MP_i = \partial m(x)/\partial x_i$ and the average product is $AP_i = x_i/m(x_i)$, the marginal product of each input can be obtained by multiplying the scale elasticities by the computed average product.

Lastly, the Allen and Morishima elasticities of substitution can be derived which are local measures of an isoquant. Substitution elasticities measure the percent changes in the factor input ratio due to a percent change in the marginal technical rate of substitution. That is the elasticity of substitution looks at how

[7] All computations were conducted using GAUSS 9.0 software.

the ratio of factor inputs changes as the slope of the isoquant changes. If a small (large) change in the slope of the isoquant results in a large (small) change in the factor input ratio, the isoquant is relatively flat (steep), resulting in a large (small) elasticity of substitution.

The Allen elasticity of substitution is defined as:

$$\sigma_{jk}^{a} = \frac{\sum_{q=1}^{Q} m_q x_q}{x_j x_k} \cdot \frac{BH_{jk}}{|BH|} \tag{19}$$

where x_q is the q^{th} input, m_q is the q^{th} partial derivative of the production function and $q = 1, ..., Q$ is the total number of inputs. The bordered Hessian is given by:

$$\begin{bmatrix} 0 & m_1 & \cdots & m_q \\ m_1 & m_{11} & \cdots & m_{1Q} \\ \vdots & \vdots & \ddots & \vdots \\ m_q & m_{Q1} & \cdots & m_{QQ} \end{bmatrix} \tag{20}$$

where the determinant of the bordered Hessian is denoted by $|BH|$ and the co-factor associated with m_{jk} in BH is denoted by BH_{jk}. The first-order, second-order, and cross-partial derivatives of the production function are denoted by m_j, m_{jj}, and m_{jk}, respectively.

An alternative elasticity is the Morishima elasticity of substitution defined as:

$$\sigma_{jk}^{m} = \frac{m_k}{x_j} \cdot \frac{BH_{jk}}{|BH|} - \frac{m_k}{x_k} \cdot \frac{BH_{jk}}{|BH|} = \frac{m_k x_k}{m_j x_j} \cdot \left(\sigma_{jk}^{a} - \sigma_{kk}^{a}\right) \tag{21}$$

which is clearly related to the Allen elasticity of substitution, but has the unique property of being asymmetric. Also, the Morishima elasticity of substitution may define two inputs as substitutes while the Allen elasticity of substitution may define that as complements. However, inputs that are Allen substitutes are always Morishima substitutes.

4 Estimation Results

The data used are from the Farming Practices Survey conducted by the Economic Research Service of the United States Department of Agriculture in 1995. The survey examined the chemical usage of farmers including nitrogen, phosphorous, and potash and the observed level of corn yield. In this study, analysis is restricted to those farmers using conventional tillage and includes 241 field level observations.[8] Summary statistics on the data are provided in Table 1. Corn yields, as well as nitrogen (N) and phosphorous (P), vary widely in the sample, and so exhibit a large standard deviation compared with the respective mean. Potash (K_2O), on the other hand, exhibits far less fluctuation in values than corn yields or nitrogen and phosphorous inputs.

Figure 1, depicts the densities of the standardized variables.[9] As can be seen from Figure 1, the densities are highly irregular, even in their standardized form. The density for corn yield and phosphorous are especially skewed, as are nitrogen and potash, though to a lesser extent. Also of note is the multimodal nature of the densities, especially for yield, nitrogen, and phosphorous.

In order to examine the application of nonparametric regression techniques to obtaining economic measures of production technology, we apply the nonparametric kernel estimator described in Section 3 to corn production in Illinois and Indiana. However, first the data is tested for known parametric forms using the J_n statistic from the Hsiao et al. (2007) consistent model specification test. The null hypothesis is: $H_o: \ E[y|x] = m(x, \beta)$, where $m(x, \beta)$ is an assumed linear or nonlinear parametric function of the independent variable matrix x and the q-dimensional parameter vector β. Specifically, the Cobb-Douglas and Translog forms in Equation 2 and Equation 3, respectively, are tested. The null hypothesis is that the model is correctly specified with rejection of the null indicating that the model indicating is incorrectly specified.

The computed test statistic for the Cobb-Douglas model was 1.698 with a marginal significance level of 0.01, while and the computed test statistic for the Translog model was 3.348 with a marginal significance level of 0.0004. Therefore, the null hypothesis of correct parametric specification for either the Cobb-Douglas or Translog is rejected at any conventional level of significance. Since the parametric forms are rejected, the nonparametric estimation procedure outlined in Section 3 is conducted using the multivariate Gaussian standard normal kernel with optimal bandwidths obtained using the LSCV method discussed above (which was found to be 0.748 for the standardized data).

Table 2 presents several calculated statistics to compare the estimated fits

[8] Additional data restrictions include omitting observations where either zero output or zero input levels have been recorded.

[9] The data are standardized by dividing by the standard deviation, and then the natural log is applied.

between the nonparametric model and the parametric counter-models. The R^2 for the Cobb-Douglas regression is 0.090 and for the Translog regression is 0.129, which is quite low compared to the nonparametric regression R^2 of 0.904. Both the mean absolute error and root mean square error for the parametric regressions are over sixty percent larger than the nonparametric regression. Finally, Theil's U_Δ statistic is computed, which gives an indication of how well the model tracks changes in the predicted observations, with lower values of the statistic indicating better forecasting ability (Greene 2005). Again, the nonparametric regression performs better than the parametric regressions, although modestly for the U_Δ statistic. In summary, the various goodness-of-fit measures computed support the Hsiao et al. (2007) specification test that the nonparametric model is preferred over the either the Cobb-Douglas or Translog models.

Next, the response coefficients, or marginal effects (i.e., the first derivatives) are estimated using Equation 9 for the nonparametric regression. The parametric regression output consists of vectors of estimated parameters, from which the estimated gradients are constructed by appropriate derivatives of the given functional form. In the current context, the estimated gradients have the interpretation of output elasticities, which measures the percent change in corn yield given a one percent change in a given fertilizer.

The estimated elasticities of output are presented in Table 3. The output elasticties vary widely between the models, though are generally all less than one and statistically significant. The Translog regression model has a negative output elasticity for phosphorous, implying that an increase in phosphorous would actually decline corn yields, however the estimate is not statistically significant. Still, the presence of a negative output elasticity in the Translog specification would seem to counter the notion of a concave production function.

Also of interest is the elasticity of scale, defined as the sum of the output elasticities, which measures the percent change in corn yield given a one percent change in all fertilizers. The production technology for Cobb-Douglas and nonparametric models would appear to exhibit decreasing returns to scale since the estimated scale elasticities are all less than one. The Translog model however indicates increasing returns to scale, though the elasticity is not significant. A hypothesis of constant returns to scale is conducted on the scale elasticities in Table 3, with the null hypothesis unable to be rejected in all of models. The hypothesis of constant returns to scale is an important test as it implies that input proportions are independent of the scale of production (Moschini 1990). In other words, constant returns to scale implies that the production function is homogeneous of degree one.

As mentioned, a benefit of nonparametric estimation is the ability to estimate values of the derivatives at every point in the sample. To shed more light on the mean computed scale and output elasticties, the nonparametric derivatives are estimated at 2.5% percent quantiles and the output and scale elasticities calculated at the given quantile. Table 7 lists the elasticities at selected quantiles to give a more global view of the agricultural technology. The output elasticities for nitrogen and phosphorous start out as negative and then increases to a positive number peaking at around the 0.50 percentile and then begins to decline. The potash output elasticity starts out positive, but also begins to increase, peaking at around the 0.55 percentile and then begins to decline. The scale elasticity tells a similar story, increasing over early quantile ranges, peaking at between 0.50 and 0.65, approaching unity, and then declining.

Table 4 reports the average and marginal product of each fertilizer for each model. Each input is associated with diminishing marginal productivity, as the average products are all generally greater than the marginal products, regardless of the regression model. This would seem to suggest that the corn farmers observed in the data set have moved beyond the optimal production allocation decision. The fact that the marginal products are less than the estimated average products would also suggest over-utilization of fertilizer input. This has important implication for input demand and imply that the representative farm could improve productivity and reduce input costs by using less fertilizer.

The marginal products are next computed using the quantile derivatives discussed above to give a more whole view of input productivity, and are plotted in Figure 2. Clearly, the productivity of fertilizer peaks very early in the production technology. Nitrogen rapidly displays rapidly increasing marginal productivity, then displays a long flat surface, and then begins to decline. Phosphorous behaves similarly, but doesn't seem to exhibit as a pronounced increase in marginal productivity as nitrogen, nor does it display a marked decline in marginal productivity. Potash seems to be mostly flat, exhibiting a spike in marginal productivity, but then quickly displays decreasing marginal productivity, remaining flat for time, then increasing, and then finally displaying decreasing marginal productivity at final production range.

The calculation of the Allen and Morishima elasticities of substitution for the Translog and nonparametric model find a clear problem in the Translog specification, namely that the Translog regression implies a non-concave production function. The Allen elasticity matrix should be negative semi-definite if farmers are assumed to maximize output amongst a set of inputs. However, the diagonal elements of the Translog Allen elasticity matrix however is not uniformly negative, thus violating concavity. However the condition is satisfied by the nonparametric regression.

The positive own-price elasticity for nitrogen in the Translog Allen elasticity matrix violates the curvature conditions necessary in order to satisfy theory. Since the input requirement set must be convex in order for the production function to exhibit quasi-concavity, a positive own substitution elasticity would imply a non-convex input requirement set, again violating production theory. The concavity problem for the Translog model is further brought to light when looking at the Morishima elasticities which take on absurdly large values. Although not reported, the hessian matrix for the Translog model is positive, providing further evidence of concavity violation, while the nonparametric hessian has diagonal negative elements and a negative determinant.

5 Implications and Conclusions

Recent technological changes have refocused attention on the estimation of primal production functions. Technologies such as precision agriculture and biotechnology raise questions that cannot be addressed using dual cost or profit functions. Given this resurgence in interest about the production function, the possible use of nonparametric regressions for the depiction of the production surface is examined. Using production data for corn from Illinois and Indiana, a multivariate production surfaces is estimated using two common parametric models, the Cobb-Douglas and Translog. The parametric specifications are then compared against a nonparametric kernel regression based on a Gaussian kernel. In terms of prediction and model fit, both the Cobb-Douglas and Translog specifications are rejected in favor of the nonparametric form.

The paper demonstrates how production surface obtained from the nonparametric kernel regression can be used to derive several measures of economic technology. The representative farm is found to exhibit constant returns to scale, a requisite for homogeneity of degree one of the production function. However, the Translog production specification is found to yield fragile results, violating concavity of the production function. Specifically, the hessian matrix is not negative semi-definite and positive own substitution elasticities are found. The nonparametric results are however consistent with concavity and do not violate theoretical conditions.

Generally, the results imply that the representative corn farm is operating at constant returns to scale, but display diminishing marginal productivity. This fact is reflected not only in the fact that factor inputs have marginal products less than one, but that the marginal products are all less than the average product. Farmers seem to over use fertilizers, especially nitrogen, which display diminishing marginal productivity along a wide range of production. Not only is this costly to the farmer in terms of lower yields, but also in terms of fertilizer cost. The results demonstrate that farmers could improve productivity and corn yield by using less fertilizer. This result is of also of importance in terms of environmental quality, as fertilizer run-off, often a result of over fertilization, causing ecological and environmental damage.

The nonparametric procedure has two key advantages over parametric specifications. First, it yields a smooth distribution function that better represents the true distribution from which corn yield is drawn, which is confirmed by the summary fit statistics discussed earlier, as well as the Hsiao-Li-Racine specification test. Second, which follows from the first reason, the nonparametric procedures offers a truer picture of the agricultural technology, which is confirmed by the satisfaction of concavity, of which the Translog specification fails.

References

[1] Ackello-Ogutu, C., Paris, Q., and Williams, W.A. (1985). Testing a von Liebig crop response function against polynomial specifications. *American Journal of Agricultural Economics*, 67(4), 873-880.

[2] Babcock, B.A. (1992). The effects of uncertainty on optimal nitrogen applications. *Review of Agricultural Economics*, 14(2), 271-280.

[3] Babcock, B.A., and Blackmer, A.M. (1994). The "ex-post" relationship between growing conditions and optimal fertilizer levels. *Review of Agricultural Economics*, 16(3), 353-362.

[4] Berck, P. and Helfand, G. (1990). Reconciling the von Liebig and differentiable crop production functions. *American Journal of Agricultural Economics*, 72(4), 985-996.

[5] Berck, P., Geoghegan, J., and Stohs, S. (2000). A strong test of the von Liebig hypothesis. *American Journal of Agricultural Economics*, 82(4), 948-955.

[6] Cooper, J.C. (2000). Nonparametric and semi-parametric recreational demand analysis. *American Journal of Agricultural Economics*, 82(2), 451-462.

[7] Chambers, R.G. and Lichtenberg, E. (1996). A nonparametric approach to the von Liebig-Paris technology. *American Journal of Agricultural Economics*, 78(2), 373-386.

[8] DiNardo, J., and Tobias, J.L. (2001). Nonparametric density and regression estimation. *The Journal of Economic Perspectives*, 15(4), 11-28.

[9] Efron, B. and R. J. Tibshirani. (1993). *An Introduction to the Bootstrap.* New York, NY: Chapman & Hall.

[10] Elbadawi, I., Gallant, A.R., and Souza, G. (1983). An elasticity can be estimated without a priori knowledge of functional form. *Econometrica*, 51(6), 1731-1751.

[11] Frank, M.D., Beattie, B.R., and Embleton, M.E. (1990). A comparison of alternative crop response models. *American Journal of Agricultural Economics*, 72(3), 597-603.

[12] Hardle, W. (1990). *Applied Nonparametric Regression.* New York, NY: Cambridge University Press.

[13] Heady, E. O. (1952). *Economics of Agricultural Production and Resource Use.* New York, NY: Prentice-Hall, Inc.

[14] Heady, E. O. (1961). *Agricultural Production Functions.* Ames, Iowa: Iowa State University Press.

[15] Henderson D.J., and Kumbhakar, S.C. (2006). Public and private capital productivity puzzle: A nonparametric approach. *Southern Economic Journal*, 73, 219-232.

[16] Henderson, D.J., and Millimet, D.L. (2007). Pollution abatement costs and foreign direct investment inflows U.S. states: A nonparametric reassessment. *Review of Economics and Statistics*, 89, 178-183.

[17] Holloway, G. and Paris, Q. (2002). Production efficiency in the von Liebig model. *American Journal of Agricultural Economics*, 84(5), 2002.

[18] Hsiao, C., Q. Li, and J.S. Racine. (2007). A consistent model specification test with mixed discrete and continuous data. *Journal of Econometrics* 140, 802: 826.

[19] Ker, A.P., and Goodwin, B.K. (2000). Nonparametric estimation of crop response insurance rates revisited. *American Journal of Agricultural Economics*, 82(2), 463-478.

[20] Lanzer, E.A. and Paris, Q. (1981). A new analytical framework for the fertilization problem. *American Journal of Agricultural Economics*, 63(1), 93-103.

[21] Li, Q., and J.S. Racine. (2007). *Nonparametric Econometrics.* Princeton, NJ: Princeton University Press.

[22] Moschini, G. (1990). Nonparametric and semiparametric estimation: An analysis of multiproduct returns to scale. *American Journal of Agricultural Economics*, 72(3), 589-596.

[23] Moss, C.B., and Schmitz, T.G. (2006). A semiparametric estimator of the Zellner production function for corn: Fitting the univariate primal. *Applied Economics Letters*, 13(20), 863-867.

[24] Mundlak, Y. (1996). Production function estimation: Reviving the primal. *Econometrica*, 64(2), 431-438.

[25] Pagan, A., and A. Ullah. (1999). *Nonparametric Econometrics.* New York, NY: Cambridge University Press.

[26] Paris, Q. (1992). The von Liebig Hypothesis. *American Journal of Agricultural Economics*, 74(4), 1019-1028.

[27] Paris, Q. and Knapp, K. (1989). Estimation of von Liebig response Functions. *American Journal of Agricultural Economics*, 71(1), 178-186.

[28] Rilstone, P. (1990). Nonparametric approach to general hypothesis testing. *Journal of Quantitative Economics*, 6(2), 331-350.

[29] Rilstone, P. (1991). Nonparametric hypothesis testing with parametric rates of convergence. *International Economic Review*, 32(1): 209-227.

[30] Rilstone, P., and A. Ullah (1989). Nonparametric estimation of response coefficients. *Communications in Statistics - Theory and Methods*, 18(7), 2615-2627.

[31] Shephard, R. W. (1970). *Theory of Cost and Production Functions*. Princeton, NJ: Princeton University Press.

[32] Ullah, A. (1988a). Nonparametric estimation and hypothesis testing in econometric models. *Empirical Economics*, 13, 223-249.

[33] Ullah, A. (1988b). Nonparametric estimation of econometric functionals. *Canadian Journal of Economics*, 21(3), 625-658.

[34] Ullah, A., and H.D. Vinod. (1988). Nonparametric kernel estimation of econometric parameters. *Journal of Quantitative Economics*, 4(1), 81-87.

[35] Vinod, H.D. and Ullah, A. (1988). Flexible production function estimation by nonparametric kernel estimators. In *Advances in Econometrics*, Volume 7, pages 139-160. Cambridge University Press, Cambridge.

[36] Wallach, D. and Loisel, P. (1994). Effect of parameter estimation of fertilizer optimization. *Applied Statistics*, 43(4), 641-651.

[37] Yatchew, A. (1998). Nonparametric regression techniques in economics. *Journal of Economic Literature*, 36(2), 669-721.

[38] Yatchew, A. (2000). Scale economies in electricity distribution: A semiparametric analysis. *Journal of Applied Econometrics*, 15(2), 187-210.

[39] Zhengfei, G., Lansink, A.O., van Ittersum, M., and Wossink, A. (2006). Integrating agronomic principles into production function specification: A dichotomy of growth inputs and facilitating inputs. *American Journal of Agricultural Economics*, 88(1), 203-214.

Table 1: Descriptive statistics

	Max	Mean	Min	Std. Dev.
Corn Yield (bu./acre)	347.90	144.26	6.00	64.04
Nitrogen (lbs./acre)	480.40	72.26	5.10	48.92
Phosphorous (lbs./acre)	360.00	110.15	4.40	51.66
Potash (lbs./acre)	224.80	121.28	41.00	25.86

Table 2: Summary of estimated fit

	Cobb-Douglas	Translog	Nonparametric
R^2	0.090	0.129	0.904
MAE	194.883	191.457	111.661
RMSE	15.288	14.886	9.983
Theil's U_Δ	0.964	1.034	0.940

Table 3: Output elasticity estimates

	Cobb-Douglas[10]	Translog	Nonparametric
Nitrogen (ϵ_1)	0.089 (0.075)	0.772 (0.389)	0.240 (0.068)
Phosphorous (ϵ_2)	0.186 (0.070)	-0.710 (0.248)	0.557 (0.068)
Potash (ϵ_3)	0.366 (0.178)	1.041 (2.039)	0.102 (0.156)
Scale Elasticity ($\sum_i \epsilon_i$)	0.641 (0.205)	1.104 (3.766)	0.899 (0.183)
t-Ratio ($H_0 : \epsilon = 1$)	-1.751	0.028	-0.552

Table 4: Estimated average and marginal products

	Cobb-Douglas[11]		Translog		Nonparametric	
	AP	MP	AP	MP	AP	MP
Nitrogen	0.779	0.069 (0.075)	0.798	0.616 (0.389)	0.424	0.102 (0.068)
Phosphorous	1.127	0.209 (0.70)	1.085	-0.770 (0.248)	1.058	0.589 (0.068)
Potash	2.374	0.869 (0.178)	2.287	2.381 (2.039)	2.250	0.230 (0.156)

Table 5: Estimated mean Allen elasticities of substitution

	Translog			Nonparametric		
	σ^a_{i1}	σ^a_{i2}	σ^a_{i3}	σ^a_{i1}	σ^a_{i2}	σ^a_{i3}
σ^a_{1i}	6.101	-2.985	-0.499	-1.725	-0.502	0.035
σ^a_{2i}	-2.985	-0.488	0.848	-0.502	-0.176	-0.066
σ^a_{3i}	-0.499	0.848	-0.146	0.035	-0.066	-0.189

Table 6: Estimated mean Morishima elasticities of substitution

	Translog			Nonparametric		
	σ^m_{i1}	σ^m_{i2}	σ^m_{i3}	σ^m_{i1}	σ^m_{i2}	σ^m_{i3}
σ^m_{1i}	0	9.833	-9.456	0	-5.930	3.390
σ^m_{2i}	22.890	0	-30.734	0.665	0	0.479
σ^m_{3i}	-11.826	-4.320	0	5.402	2.840	0

Table 7: Nonparametric output and scale elasticities at given quantiles

Quantile	Nitrogen (ϵ_1)	Phosphorous (ϵ_2)	Potash (ϵ_3)	Scale Elasticity ($\sum_i \epsilon_i$)
0.025	-0.286	-0.191	0.055	-0.422
0.050	-0.258	-0.121	0.106	-0.273
0.075	-0.134	0.113	0.161	0.141
0.100	0.042	0.396	0.117	0.555
0.475	0.232	0.481	0.089	0.802
0.500	0.234	0.485	0.097	0.816
0.525	0.235	0.487	0.103	0.825
0.550	0.236	0.488	0.105	0.829
0.900	0.185	0.305	0.088	0.578
0.925	0.185	0.302	0.086	0.572
0.950	0.179	0.286	0.077	0.542
0.975	0.157	0.244	0.052	0.453

Figure 1: Density Estimates for Corn Yield and Fertilizer Input

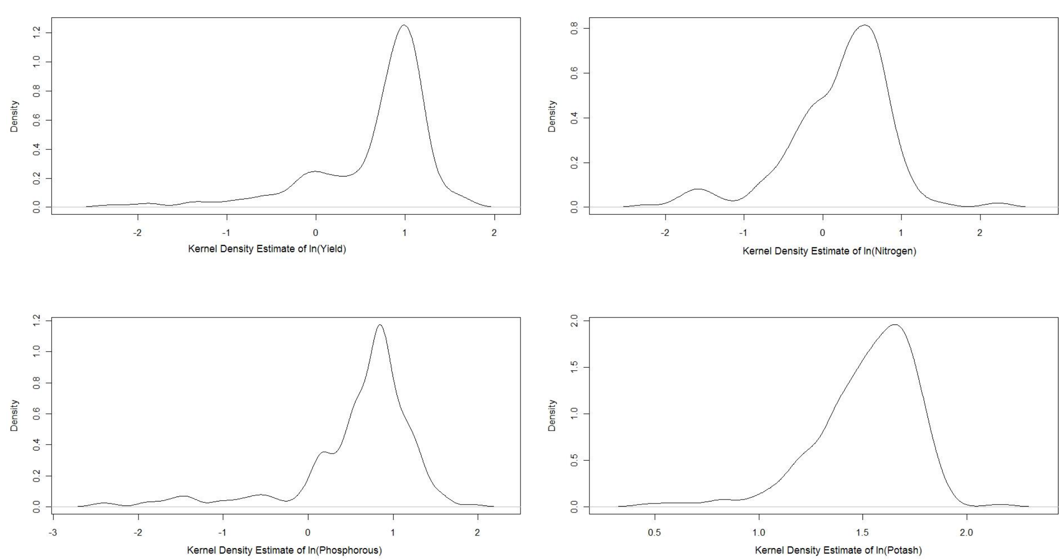

Figure 2: Marginal products of fertilizer inputs at given quantiles