Invited Presentation

A Brief Overview of Nonparametric Methods in Economics

Arne Hallam

The concept of nonparametric analysis, estimation, and inference has a long and storied existence in the annals of economic measurement. At least four rather distinct types of analysis are lumped under the broad heading of nonparametrics. The oldest, and perhaps most common, is that associated with distribution-free methods and order statistics. Similar in spirit, but different in emphasis, is nonparametric density estimation, such as the currently popular kernel estimator for regression. Semi-parametric or semi-nonparametric estimation combines parametric analysis of portions of the problem with nonparametric specification for the remainder, such as the specification of a specific functional form for a regression function with a nonparametric representation of the error distribution. The final type of nonparametrics is that associated with data envelopment analysis and revealed preference, although the use of the term nonparametrics for this research is perhaps a misnomer. This paper will briefly review each of the four types of analysis, leaning heavily on other published work for more detailed exposition. The paper will then discuss in more detail the application of the revealed-preference approach to four specific economic problems: efficiency, the structure of technology or preferences, technical or taste change, and risky choice. The paper is not complete, exhaustive, or detailed. The primary purpose is to expose the reader to a variety of techniques and provide ample reference to the relevant literature.

Nonparametric Inference

Much of nonparametric inference is based on ranks and order statistics. Early papers include those by Hotelling and Pabst, Friedman, and Kendall (1938). The big boost came with the 1945 paper of Wilcoxin, followed by an important theoretical paper of Mann and Whitney, and the classic monograph by Kendall (1948). The current state of the practice is represented well by the textbooks of Lehmann, Randles and Wolfe, and Gibbons. A couple of simple examples will give the flavor of the analysis.

Much of nonparametrics is based on the use of order statistics. Consider a random sample $X_1, X_2, \ldots, X_n$ from a population with continuous cumulative distribution function F_X. Now let $X_{(1)}$ denote the smallest element of the sample, and rearrange the sample in increasing order of magnitude. The set of variables $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ are termed the order statistics of the sample, and $X_{(r)}$ is called the rth order statistic. Using the probability-integral transformation, it can be shown that the variables $F(X_{(1)}) \leq \ldots \leq F(X_{(n)})$ are distributed as the order statistics from a uniform distribution on (0,1) (Randles and Wolfe, p. 7). Thus, no matter what the underlying distribution of X, these statistics are uniformly distributed. In this sense, they are referred to as distribution-free. Joint and marginal distributions, moments, and asymptotic properties of these statistics are easily found (Gibbons, pp. 24–29), and analysis can proceed without troublesome assumptions about the nature of the underlying distributions.

In a similar manner, analysis can be based on the ranks of the variables in this ordering. For example consider two populations, one of which is subjected to a treatment. For example, consider the starting salaries of new professors with training in economics as opposed to finance. Rank all the new professors by starting salary, with the random variable for each individual being his or her rank in the sample. The Wilcoxin rank-sum test is performed by adding the ranks of all those trained in economics and comparing it to some critical value. If this sum is sufficiently large, then one can reject the hypothesis that those trained in finance have higher salaries.

While nonparametric methods avoid many problems associated with specification error, they are generally less efficient than parametric methods if the form of the underlying distribution is known.

Arne Hallam is an associate professor, Department of Economics, Iowa State University.

Journal Paper no. J-15075 of the Iowa Agricultural and Home Economics Experiment Station, Ames, IA. Project no. 2894.

While the true form of underlying distributions is never known, extensive Monte-Carlo work has been performed comparing order and rank statistics to parametric alternatives for many standard distributions (Lehmann). While the results support parametric alternatives when the distribution is known, they also argue for specific Monte Carlo work for individual problems and situations.

Given economists' preoccupation with model building and regression analysis, these techniques have not had heavy use. There has been increasing interest in these techniques in recent years, however. Holmes and Hutton have proposed the multiple rank F test for investigating causality between economic variables as an alternative to the standard Granger tests. Pagan and Schwert propose several nonparametric tests for covariance stationarity and use them in analyzing stock-market data. Such techniques could be a useful precursor to standard parametric time series analysis. Campbell and Dufour use sign statistics and Wilcoxin signed-rank statistics to test the independence of time series that may be subject to the Mankiw-Shapiro criticism due to feedback and rational expectations. They found that these nonparametric tests avoid the too frequent rejection problem associated with standard parametric tests. While more closely related to density estimation discussed in the next section, Han (1987a, b) and Matzkin (1991a) have extended the concept of rank correlation due to Kendall (1938) to develop a maximum rank correlation estimator that generalizes in a nonparametric way the Box-Cox transformation. They analyze nonparametric transformations of the linear model of the following form:

(1) $y_i = D \cdot F(x_i'\beta_0, \epsilon_i \quad i = 1, \ldots, n),$

where the composite transformation $D \cdot F$ is such that D: $R \rightarrow R$ is nondegenerate monotonic and F: $R^2 \rightarrow R$ is strictly monotonic in its arguments. While $x'\beta_0$ is still a linear function, it can affect y in a nonlinear and nonparameterized fashion that is not separable in the error term. Such models allow for more general transformations of the variables, but at the cost of computational complexity as the estimation usually involves a grid search. The greatest value of these papers is perhaps in reminding us that old ideas can often be revived and modified to solve new problems.

The use of standard nonparametric techniques is due for a resurgence in economics. These techniques are particularly useful for comparing the properties of two populations. Comparing yield distributions over time and across space is one possible application. The techniques are also useful for analyzing the residuals from estimated models such as the standard K-S test for normality. They could also prove useful in investigating the properties of random variables generated using optimization techniques or parameters estimated by simulation. An example is the work of Tolley and Pope on stochastic dominance tests.

Nonparametric Regression Models

Two approaches to nonparametric curve fitting have become popular in the last decade. The first is the use of series approximations popularized by Gallant (Gallant 1981, 1982; Gallant and Nychka; Andrews). Spline approximations would also fit in this category (Engle et al.). The second is kernel estimation, or smoothing, as discussed by Bierens, Ullah, and in the recent monograph by Hardle. White and Wooldridge argue that both of these techniques are special cases of the method of sieves proposed by Grenander. These and other related techniques are discussed in the recent survey paper of Pagan and Wickens.

Consider first the advantages of series estimators. The economist is often interested in estimating the moments of a variable y conditional on the realization of some random vector $x = \bar{x}$. Denote this conditional moment as $\theta(x_t) = E(y_t|x_t = \bar{x}_t)$. The regression model is then written as

(2) $$y_t = m(x_t) + \epsilon_t,$$

where $y_t, \epsilon_t \in R$, $x_t \in D \subset R^d$, $m(\cdot) \in M$, and M is the class of continuous functions from D to R. The set D is the set of d-element vectors considered as the sample space for the variable x. The objective is to estimate various functions of $m(\cdot)$, such as $m(\bar{x})$ and the derivatives of m for arbitrary values of x. A parametric approach would then approximate this unknown function with a specific functional form, such as a quadratic, translog, or CES. Significant work has been done on the various advantages and disadvantages of various functional forms (Barnett and Lee; Diewert and Wales). One difficulty is that these approximations may not be globally valid. Another is that they do not improve as the sample size increases. The nonparametric approach is to approximate $m(\cdot)$ by a finite series expansion of the following form:

(3) $$m(\cdot) \approx \sum_{s=1}^{k_n} z_s(\cdot)\gamma_s,$$

where $\{z_s: s = 1, 2, \ldots\}$ is a family of functions (such as the trigonometric ones) from D to R, $\gamma = (\gamma_1, \ldots, \gamma_{kn})$ is an unknown parameter vector, and k_n is the number of summands in the series expansion when the sample size is n. For example, Gallant proposed that z include linear and quadratic terms in x plus the trigonometric terms $\cos(b'x)$ and $\sin(b'x)$ for some integer vector $b \in R^d$. The number of terms included in the trigonometric series depends explicitly on the sample size, and thus both the number of parameters and, hopefully, accuracy increase as the sample size increases. Andrews has given conditions on the rate at which the number of terms in the series must approach infinity, as the sample size increases to ensure consistency and asymptotic normality of the estimator. These series estimators can be also be viewed as a type of GMM estimator, where the orthogonality condition holds in the limit as the number of terms in the series increases (Pagan and Wickens). Since these estimators almost always include parametric as well as series-expansion terms and parameterize the series expansion, they are often called semi-nonparametric. The primary advantage of this type of estimator is its ability to approximate any underlying functional relationship to a high degree of accuracy as the sample size increases. Disadvantages include a possibly large number of parameters and the distinctly periodic behavior of the estimated relationship for series with small numbers of terms (Pope). For a discussion of these and other issues, see the symposium papers contained in the 1984 *American Journal of Agricultural Economics* (Gallant 1984). Recent papers using this approach include Chalfant, Pagan and Hong, and Barnett and Yue.

Kernel estimation is based on the work of Rosenblatt, Nadaraya (1964a, b), and Watson. Two standard references are Prakasa Rao and Silverman. The use of kernel estimators for time series is discussed by Robinson (1983), while a recent text with economic examples is Hardle. The kernel estimator is basically a method of smoothing. A smoothing for $m(x)$ in (2) is given by a weighted average of values of y_t in the neighborhood of x_t. Specifically, for the case of a single x variable, a smoothing estimator would be

$$(4)\qquad \hat{m}(x) = \frac{\sum_{t=1}^{n} W_{nt}(x)y_t}{n},$$

where W_{nt} is a weight that may depend on the entire set of x variables.

As the kernel estimator is a smoothing estimator of the conditional expectation of y in (2), consider the joint density of y and a $k \times 1$ x vector. Because $m(\cdot)$ is the expected value of y given x, its value is

$$(5)\qquad m(x) = \int y f(y|x)\, dy = \int y \frac{f(y,x)}{f(x)}\, dy, \quad \text{if} \quad f(x) > 0,$$

where $f(y|x)$ is the conditional density of y given x, $f(x,y)$ is the joint density, and $f(x) = \int f(y,x)\, dy$ is the marginal density of x. The function m can be estimated by obtaining estimates of $f(y,x)$ and $f(x)$. The kernel estimator of $f(x)$ is given as

$$(6)\qquad \hat{f}_n(x) = \left(\frac{1}{n}\right) \sum_{t=1}^{n} \frac{K\left[\frac{x - x_t}{h_n}\right]}{h_n^k},$$

where $K(\cdot)$ is a chosen real function on R^k that satisfies

$$(7)\qquad \int |K(x)|\, dx < \infty \quad \int K(x)\, dx = 1.$$

The parameter h is called a window width parameter and satisfies

$$(8)\qquad \lim_{n\to\infty} h_n = 0 \quad \lim_{n\to\infty} nh_n^k = \infty.$$

A kernel estimator of $f(y,x)$ consistent with the derived estimator of the marginal density $f(x)$ is

$$(9)\qquad \hat{f}(y,x) = \left(\frac{1}{n}\right) \sum_{t=1}^{n} \frac{K^*\left[\frac{(y - y_t)}{h_n}, \frac{(x - x_t)}{h_n}\right]}{h_n^{k+1}}$$

where K^* satisfies

$$(10)\qquad \int yK^*(y,x)\, dy = 0 \quad \int K^*(y,x)\, dy = K(x).$$

Clearly, $\hat{f}(x)$ is the marginal density since

$$(11)\qquad \hat{f}(x) = \int \hat{f}(y,x)\, dy$$

$$= \int \left(\frac{1}{n}\right) \sum_{t=1}^{n} \frac{K^*\left[\frac{(y - y_t)}{h_n}, \frac{(x - x_t)}{h_n}\right]}{h_n^{k+1}}\, dy$$

$$= \left(\frac{1}{n}\right) \sum_{t=1}^{n} \int \frac{K^*\left[\frac{(y - y_t)}{h_n}, \frac{(x - x_t)}{h_n}\right]}{h_n^{k}}\, d\left(\frac{y}{h_n}\right)$$

$$= \left(\frac{1}{n}\right) \sum_{t=1}^{n} \frac{K^*\left[\frac{(x - x_t)}{h_n}\right]}{h_n^{k}}$$

The estimator of $m(x)$ is then given by

(12)

$$\hat{m}(x) = \frac{\int y\hat{f}(y,x)\,dy}{\hat{f}(x)} = \frac{\int y \left(\frac{1}{n}\right) \sum_{t=1}^{n} \frac{K^*\left[\frac{(y-y_t)}{h_n}, \frac{(x-x_t)}{h_n}\right]}{h_n^{k+1}}\,dy}{\hat{f}(x)} = \frac{\left(\frac{1}{n}\right) \sum_{t=1}^{n} y_t \frac{K\left[\frac{(x-x_t)}{h_n}\right]}{h_n^k}}{\hat{f}(x)} = \frac{\sum_{t=1}^{n} y_t K\left[\frac{(x-x_t)}{h_n}\right]}{\sum_{t=1}^{n} K\left[\frac{(x-x_t)}{h_n}\right]}.$$

This estimator is a weighted mean of the dependent variable y_t. The closer x is to x_t, the more weight is put on y_t. There is a large literature on the choice of the appropriate kernel. Common choices include the standard normal ($K(x) = (2\pi^{-1/2})\exp(-\frac{1}{2}x^2)$) and various polynomials. The properties of the estimators are not particularly sensitive to the choice of the kernel function. Bierens, Epanechinikov, and Hardle (chapter 4) each contain a more detailed discussion of this issue. Of perhaps more serious concern is the choice of the bandwidth, h. The bandwidth is chosen so that its rate of decrease with the sample size will give desirable asymptotic properties for the estimator. If h is chosen to be too large, the estimate will be too smooth and obscure the shape of the density function. If h is too small, the estimate of $m(x)$ will follow the data closely and may result in overfitting. The literature suggests several way to choose h, most giving forms similar to $h = cn^{-b}$, where b is a fraction. While several suggestions are given for choosing h in an ''optimal'' fashion, in practice the choice has tended to be arbitrary.

There has been an explosion of papers using kernel estimators. Some of these combine parametric with kernel specifications. Vinod and Ullah consider the estimation of production functions using this technique. Moschini combines parametric and nonparametric kernel estimation in investigating preference change in meat demand. Stoker uses a kernel estimator to test additive constraints on the first and second derivatives of a regression model. Pagan and Ullah discuss the use of kernel estimators for incorporating risk as a regressor. Hardle, Hildenbrand, and Jerison use a nonparametric kernel estimator to estimate the mean income effect matrix in a cross section of demands. Rilstone examines the use of kernel estimators for average, as opposed to point, estimates of derivatives. Robinson (1991) uses kernel estimators to estimate the Kullback-Liebler information criterion in forming a test of the independence of time series data. McCurdy and Stengos compare nonparametric kernel estimators to traditional parametric estimators in studying the time pattern of the risk premium, while Hong and Pagan compare kernel and series estimators using Monte Carlo techniques.

Kernel estimators avoid many of the specification problems associated with parametric functional forms. They also have a number of limitations. Kernel estimators may be sensitive to the chosen bandwidth. To the extent that the bandwidth influences the results, there is arbitrariness in any reported estimate. Nonparametric estimators have strong asymptotic properties. They may not be particularly useful in small samples since smoothing tends to be more useful with large amounts of data. Given the sample sizes used in many studies employing annual data, kernel estimators may see limited use in agricultural economics.

Semiparametric Econometrics

Semiparametric models combine a parametric component of known form with a nonparametric or infinite dimensional one. Consider a simple linear regression model of the form

(13) $$y_t = x_t'\beta + \epsilon_t,$$

where the error ϵ_t has density function $f(\cdot)$. If the errors are identically and independently distributed, ordinary least squares estimates of β are Gauss-Markov efficient. If the errors are also Gaussian and satisfy standard regularity conditions, the variance of the least squares estimator achieves the Cramer-Rao lower bound. If the errors are heteroskedastic or serially dependent with known covariance variance, the generalized least squares (GLS) estimator is efficient; but, if this covariance matrix is not known, the GLS estimator is not defined. Feasible GLS provides a way to obtain consistent estimates of the parameters of the model but requires some assumptions about the properties of ϵ. In this model, a semiparametric estimator would be one that used a nonparametric technique to obtain estimates of the unknown covariance matrix. Robinson (1988), in his survey of semiparametric methods, suggests that semiparametric techniques might be viewed as a way to avoid parameterizing nuisance parameters in econometric models. In the above model, the properties of ϵ as regards to heteroskedasticity, serial correlation, or form of the underlying distribution (normal, log-normal, beta, etc.) are really nuisance parameters as far as obtaining estimates of β. If these properties can be parameterized (such as normality with first-order serial correlation and homoskedasticity), feasible generalized least squares (in this case, for example, use of the Cochrane Orcutt procedure) is consistent and achieves the Cramer-Rao lower bound asymptotically. Semiparametric estimation may be viewed then as the parametric estimation of (13), without parametric specification of the error structure $f(\epsilon_t)$.

Consider the log-likelihood function for estimation of (13):

$$L = \sum_{t=1}^{n} \log f(y_t - x_t'\beta). \tag{14}$$

One approach (Gallant and Nychka) would be to replace the unknown $f(\cdot)$ with the product of a $N(0,\sigma^2)$ variable denoted by $\phi(\epsilon)$ and a polynomial in ϵ, $P(\epsilon)$. The approximate log-likelihood function L^* is then given as

$$L^* = \sum_{t=1}^{n} \log \phi(y_t - x_t'\beta) + \sum_{t=1}^{n} \log P(y_t - x_t'\beta). \tag{15}$$

Gallant, Hsieh, and Tauchen have used this formulation in the study of exchange rates.

An alternative to approximation is to solve the first-order conditions that come from the scores associated with (13). The score, by definition, is the derivative of the log-likelihood function with respect to the parameters. For (14), this gives

$$-\sum_{t=1}^{n} x_t f(\epsilon_t)^{-1}\left[\frac{\partial f(\epsilon_t)}{\partial \epsilon_t}\right]. \tag{16}$$

Nonparametric estimates of the density and its derivatives could then be used to find the values of β that optimize the scores. Stone was the first to use this method, and Manski (1984, 1985, 1987) provides several extensions. A recent paper by Horowitz uses the kernel estimator to smooth the scores for a binary response model. Lee proposes a nonlinear least squares estimator similar to a kernel estimator for the truncated regression model. Several of the suggested approaches involve OLS estimation and then use of the residuals as estimates of the underlying empirical density function. Given some structure on the underlying problem, semiparametric estimation may reduce to finding an optimal nonparametric estimate of some particular moment of the unknown distribution (Chamberlain 1987; Newey 1988). Pagan and Wickens give a GMM interpretation of semiparametric estimators.

An important concept in semiparametric estimation is the property of adaptation. A semiparametric estimate is said to be "adaptive" if it is as efficient over a broad class of assumptions concerning ϵ as a correct finite parameterization (Stein). Another important concept in semiparametric estimation is the efficiency bound. This is an extension to the semiparametric problem of the Cramer-Rao bound and can be viewed as representing the information lost from using a nonparametric estimator as opposed to a correctly parameterized one. This topic is currently being intensively investigated. The survey by Robinson (1988) on semiparametric methods analyzes these bounds in some detail. Coslett discusses bounds for binary choice and censored regression models. Chamberlain (1992) discusses bounds using conditional moment restrictions, while Newey (1990) provides a lucid comprehensive survey. Recent application papers using semiparametric methods are Sentana and Wadhwani using stock-market data, Deaton on income distribution in Thailand, Moschini on Ontario dairy farms, and Holt and Moschini on sow farrowing decisions.

Semiparametric estimators may have a significant role in future econometric work. They allow for flexible specification of nuisance and other unknown parameters. They are particularly useful in analyzing large data sets with small numbers of variables. They have some drawbacks, however. While they have good asymptotic properties, they may do poorly in small samples compared to incorrectly specified parametric estimators since their rate of convergence is typically slower than usual maximum-likelihood estimators. They are not really practical for large numbers of parameters or moment restrictions. Two-step implementations exploiting the parametric portions may alleviate some of these problems.

Revealed Preference, Behavior, and "Nonparametric" Analysis

The work of Farrell on efficiency and Afriat (1967, 1972, 1973) on revealed preference has given rise to another type of analysis that is sometimes called nonparametric. This research attempts to understand technology and behavior without imposing any functional form on the data. However, it differs from the nonparametric analysis previously discussed in that statistical inference is not explicitly or implicitly used. In some ways, it is a kind of reduced-form analysis. Consider, for example, a consumer who maximizes utility subject to a budget constraint. Certain assumptions about the utility function lead to restrictions on the derived demand functions. The nonparametric approach would examine the actual choices of a consumer and ask if there exists a utility function consistent with the observed choices and the assumptions. If one exists, the researcher concludes that the utility maximization hypothesis and the assumptions are valid. The answer to the question, however, can only be yes or no. The methods do not give probabilities of the hypothesis being true. These methods are nonparametric in the sense that no functional form is assumed. They are not nonparametric in the sense of nonparametric statistical analysis. Better terms for this analysis might be revealed behavior analysis (RBA) or revealed structure analysis (RSA). Revealed-behavior approaches have been used in both consumption and production analysis. The basic idea is the same, but the emphasis has typically been different. Work in production has tended to concentrate on characterization of the isoquant and efficiency, while work in consumption has emphasized tests of optimizing behavior and the structure of preferences. The discussion will proceed by first discussing the efficiency problem in some detail and then moving to more general revealed-preference arguments. The discussion of efficiency is presented as a basis for illustrating the nonparametric approach and is not a comprehensive survey.

A Digression on Efficiency Analysis

Modern efficiency analysis originated with the pioneering work of Farrell. In essence, he developed a way to measure how far a given input vector is from the boundary of the input requirement set. In order to consider this measure in more detail, consider a set of assumptions on the technology. Assume that the firm produces m outputs y using n inputs x. The input correspondence $y \rightarrow L(y) \subseteq R^n_+$ is the subset of all input vectors capable of producing at least output vector y. The correspondence is assumed to have the following properties:

(17)

$L.1 \quad 0 \notin L(y) \quad L(0) = R^n_+.$

$L.2 \quad \|y^\ell\| \rightarrow +\infty \text{ as } \ell \rightarrow +\infty \Rightarrow \bigcap_{\ell=0}^{\infty} L(y^\ell) = \varnothing.$

$L.3 \quad \text{If } x \in L(y),\ \lambda x \in L(y) \quad \text{for} \quad \lambda \geq 1.$

$L.4 \quad L$ is a closed correspondence.

$L.5 \quad L(\theta y) \subseteq L(y) \quad \theta \geq 1.$

The properties are standard and are discussed in more detail by Färe (1988), and Knox Lovell and Schmidt. Three important subsets of $L(y)$ are used to discuss efficiency. They are the isoquant and the weak-efficient subset and the efficient subset. They are defined as

$$(18) \quad \begin{aligned} Isoq\, L(y) &\equiv [x: x \in L(y), \\ &\qquad \lambda x \notin L(y),\ \lambda \in [0,1)],\ y \geq 0, \\ WEff\, L(y) &\equiv [x: x \in L(y), \\ &\qquad u <^* x \Rightarrow u \notin L(y)],\ y \geq 0, \\ Eff\, L(y) &\equiv [x: x \in L(y), \\ &\qquad u \leq x \Rightarrow u \notin L(y)],\ y \geq 0, \end{aligned}$$

where the notation $u <^* x$ implies $u_i < x_i$ or $u_i = x_i = 0$ for all i. It is obvious the $Eff\,L \subseteq WEff\,L \subseteq Isoq\,L$. Consider Figure 1 with two variable inputs. The isoquant is given by ABCD, the weak efficient set by BCD, and the efficient set by CD. Consider

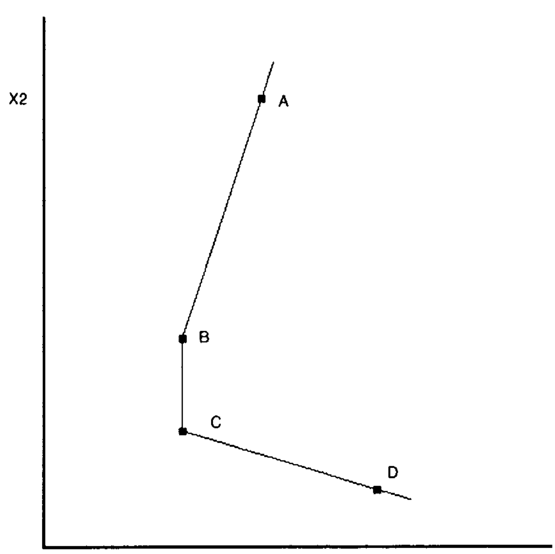

Figure 1. Efficient Subsets of $L(Y)$

now the Farrell measure of technical efficiency. It is given by

$$(19)\quad F(x;y) \equiv \min\,[\lambda\colon \lambda x \in L(y),\ \lambda \geq 0].$$

The Farrell measure computes the ratio of the smallest feasible radial contraction of an observed input vector to the input vector itself. Call this minimum value of λ, λ^0. If $\lambda^0 < 1$, the firm can produce y with a radially smaller input vector and the observed input vector is inefficient. Farrell also considers cost efficiency. Consider the cost-minimization problem for the firm:

$$(20)\qquad C(y,w) \equiv \min_x\,[w'x\colon x \in L(y)].$$

A measure of cost or overall efficiency (CE) is given by the ratio of actual cost to the minimum cost of producing output vector y:

$$(21)\qquad CE(x;y,w) \equiv \frac{C(y,w)}{w'x}.$$

Farrell defines allocative efficiency as the ratio of cost efficiency to technical efficiency or

$$(22)\qquad AE(x;y,w) \equiv \frac{CE(x;y,w)}{F(x;y)}.$$

These can all be illustrated graphically using Figure 2. Consider an input combination P. Technical efficiency is given by the ratio of OQ to QP. Cost efficiency is given by the ratio of OS to OP and allocative efficiency by the ratio OS to OQ. In a similar way, a weak input measure of technical efficiency can be defined (Färe, Grosskopf, and

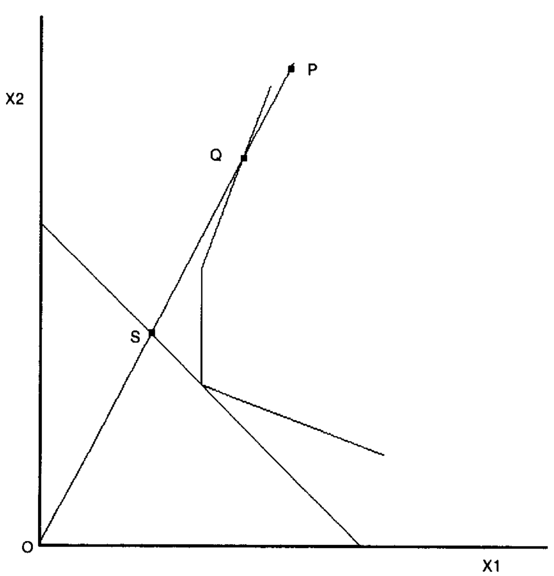

Figure 2. Farrell Efficiency

Knox Lovell). The two measures are equal only with strong disposability of inputs. A variety of nonradial measures of efficiency can also be defined (Russell). Efficiency can also discussed in terms of cost and profit functions (Kopp and Diewert).

Consider now nonparametric measurement of the technology and various efficiency measures. Let there be K firms each using n inputs to produce m outputs. Let M be a $K \times m$ matrix of the output levels, N a $K \times n$ matrix of input levels, y the output vector containing the total of each output, and x the vector of total inputs. Shephard; Färe, Grosskopf, and Knox Lovell; and Banker, Charnes, and Cooper have shown how to construct an input requirement set satisfying weak disposability of inputs that is consistent with observed data. This is done by constructing the convex weak-disposal hull of the observed data. The input set so constructed is the smallest set the includes all K observations in the sample and satisfies the properties L.1–L.5. For the case of constant returns to scale, this set is given by

$$(23)\qquad \begin{aligned} L(y) &= [x\colon z \in R^K_+,\ \mu z'M = y,\\ \delta x &= z'N,\ \mu,\delta \in (0,1)], \end{aligned}$$

where z is an intensity vector that specifies combinations of the various firms' technologies used to construct the reference technology, and μ and δ scale output and input respectively. Modifications to the set for alternative assumptions concerning returns to scale and disposability are discussed by Knox Lovell and Schmidt, as well as by Färe, Grosskopf, and Knox Lovell. The isoquant is constructed in essence by enveloping the data. If all the firm's outputs were the same, this could be visualized as finding the set of convex facets closest to the origin that allow production of the output y.

The Farrell efficiency measure is computed for each firm in the sample by solving the following programming problem.

$$(24)\quad \begin{aligned} F(x^k,y^k) &= \min_x[\lambda\colon z \in R^K_+,\ \mu z'M = y^k,\\ \lambda\delta x^k &= z'N,\ \mu,\delta \in (0,1)]\\ &= \min[\lambda\colon \lambda x^k \in L(y^k)]. \end{aligned}$$

As written, the problem is not well defined due to the scalars μ and δ, but can be written in a more standard programming format by using a change of variables. Define $s = z/\delta$, $\gamma = (1/\mu\delta)$ and then write the problem as

$$\min_{s,\gamma,\lambda} \lambda \tag{25}$$
$$\text{s.t.}$$
$$s'M - \gamma u^k = 0$$
$$s'N - \lambda x^k = 0$$
$$s \geq 0, \quad \gamma \geq 1.$$

The solution finds the shortest radial distance from the input-output combination of firm k to the isoquant constructed from the technologies of all the firms. Minimum cost can be obtained by solving the program in (20) given the technology in (23). Cost and allocative efficiency are then easily computed using these solutions.

Several points about the nonparametric approach can now be made. The approach is data enveloping in the sense that it finds the "minimal" set of points that satisfy some hypothesis. In this case, it finds the smallest input requirement set that satisfies the axioms and is consistent with the data. In a different context, it might find the set of points consistent with utility or profit maximization. In some cases, this set might be empty (no feasible solution), and the researcher would conclude that there is no set of points that is consistent with the hypothesized axioms and the data. The particular constraints that give rise to infeasibility may be useful in analyzing the problem, however. The reference technology is then used to construct various other useful measures such as efficiency and the cost-minimal input function. For example, the efficiency measures could be related to firm size or country of origin.

Revealed Preference and Nonparametric Analysis

While the work of Farrell is the precursor to the nonparametric revealed-behavior approach in production, the work of Afriat (1967, 1973) is the foundation of analysis for consumers. In his original paper, Afriat (1967) developed a set of inequalities that must be satisfied by observed price and individual demand data if they were generated by utility maximization. He termed these cyclical, multiplier, and level consistency. These basic inequalities have been slightly modified and extensively discussed (Diewert 1973a, 1978; Diewert and Parkin 1985; Varian 1982, 1983a). Similar to the construction of the isoquant in efficiency analysis, they provide a set of conditions that data generated by a utility-maximizing consumer must satisfy. In order to discuss these, three definitions are needed (Varian 1982, 1983a).

Definition 1: A utility function, $U(x)$, rationalizes the data (p^i, x^i), $i = 1, \ldots, n$, if $U(x^i) \geq U(x)$ for all x such that $p^i x^i \geq p^i x$, for $i = 1, \ldots, n$.

Definition 2: An observation x^i is directly revealed preferred to a bundle x, written $x^i R^0 x$, if $p^i x^i \geq p^i x$. An observation x^i is revealed preferred to a bundle x, written $x^i Rx$, if there is some sequence of bundles $(x^j, x^k, \ldots, x^\ell)$ such that $x^i R^0 x^j$, $x^j R^0 x^k$, $\ldots$, $x^k R^0 x^\ell$.

Definition 3: The data satisfies the Generalized Axiom of Revealed Preference (GARP) if $x^i R x^j$ implies $p^j x^j \geq p^j x^i$.

Afriat's theorem then says that the following four conditions are equivalent:

1. There exists a nonsatiated utility function that rationalizes the data.
2. The data satisfies GARP.
3. There exist numbers $\mu^i, \lambda^i > 0$, $i = 1, \ldots, n$ that satisfy the Afriat inequalities: $\mu^i \leq \mu^j + \lambda^j p^j(x^i - x^j)$ for $i,j = 1, \ldots, n$.
4. There exists a concave, monotonic, continuous nonsatiated utility function that rationalizes the data.

The point is that if there exists a set of numbers that rationalize the data, there exists a well-behaved utility function and vice versa. Tests of the utility-maximization hypothesis can then be constructed by considering whether there exists a solution to a set of linear inequalities. This problem was initially solved by Afriat using linear programming. Diewert and Parkin (1985) provide extensions of this approach. A major contribution of Varian was the development of solution techniques for some types of problems that do not require the solution of a programming problem. In fact, much of the current popularity of the approach is the ease of use of Varian's canned procedures.

While the initial paper of Afriat (1967) concentrated on the existence of a utility function rationalizing the data, more recent work (Varian 1983a; Diewert and Parkin 1985) has emphasized testing for the existence of a utility function satisfying certain properties and also rationalizing the data. Thus, the approach can be used to test for homotheticity, separability, and even technical change.

Contemporaneous with the development of the revealed-preference approach to consumer problems, work proceeded on nonparametric approaches to production analysis (Afriat 1972; Hanoch and Rothschild). Again the approach was to develop a set of conditions that must be satisfied by the data if they were generated by a producer optimizing some objective function subject to some technology. Varian (1984) again provided extensions of the original work along with suggestions for easy computation for many of the simpler tests. An excellent discussion of testing using programming techniques is Diewert and Parkin (1983). Banker and Maindiratta have extended Varian's approach to cases when all the firms may not satisfy the assumption of profit maximization and link Varian's results to those suggested by data envelopment. As in the consumer case, the tests can be used to test for specific structures as well as for consistency with optimization hypotheses and an underlying technology. The early work of Farrell is implicit in most of this research.

The revealed-behavior approach has many advantages over the traditional parametric approach to production and consumption. It allows for the "testing" of economic hypotheses independent of functional form. This is important since a separability test using a translog form is a simultaneous test for separability and the translog form. The approach also has several drawbacks. The primary one is that the tests as currently implemented have no statistical properties. The data either satisfy the axioms or they do not. Consider, for example, the case when the appropriate linear-programming problem has an infeasible solution. Suppose that the constraints (inequalities) are only violated at one point and that the violation is of a small order of magnitude. Does this mean that the data is not consistent with the maintained hypothesis, or does it mean that a very small error occurred somewhere (measurement, optimization, model structure, etc.)? Varian (1985, 1990) has addressed the measurement error issue in some detail. His approach is to consider the degree of measurement error necessary for the data to satisfy the axioms. Tsur has developed a simpler way to implement the approach for large data sets. Epstein and Yatchew have proposed an alternative way of analyzing the problem using a nonparametric approach to regression analysis. While this approach shows promise, it requires that the variance of the error term be known. Yatchew (1992) has extended the approach in ways that are less restrictive, but some prior information is still required. Work by Matzkin (1991b) on the semiparametric estimation of response subject to concavity and monotonicity constraints may also prove relevant. Bronars has proposed a way to compute the power of nonparametric tests by comparing the actual pairs with those that would occur if the data points were randomly generated. Aizcorbe has developed a lower bound based on the same approach that is not as numerically intensive. Nonparametric statistical approaches might prove to be a valuable tool in this area. Since the revealed-behavior approach considers all data points in the analysis, outliers may seriously bias the results. Whereas regression weights the various points so that a single outlier has minimal significance, this approach has no way to weight effectively the various observations.

Applications Using the Revealed-Behavior Approach

Analysis of Economic Efficiency

The work of Farrell was rapidly applied by researchers in agricultural economics. The early work was almost all associated with the University of California at Berkeley (Boles 1966; Bressler; Seitz). Most research concentrated on identification of the production frontier (efficiency locus or isoquant) and measurement of deviations from it. Boles (1971) developed linear-programming software allowing for efficient estimation of efficiency frontiers.

An early digression was the work of Aigner and Chu, and later Timmer on deterministic parametric frontiers. They avoided some of the restrictive assumptions of Farrell's work by postulating a functional form for the production function and then estimating it subject to the constraint that the residuals all be nonpositive. Thus, they moved the analysis from input space to input-output space and converted a parameterless technique of enveloping the input requirement set to the estimation of parametric production functions. This approach is probably what led to differentiating the Farrell approach as nonparametric. Once the problem was formulated as a type of least squares problem, the possibility of stochastic assumptions and statistical inference became obvious. A whole new approach and growth industry was spawned by the stochastic frontier method proposed by Aigner, Knox Lovell, and Schmidt, and simultaneously by Meeusen and van den Broeck. They proposed estimating a production function with a composed error consisting of a symmetric part reflecting measurement error and a one-sided part reflecting inefficiency on the part of producers not on the frontier. This approach dominated efficiency measurement for many years. Excellent surveys are Schmidt, Knox Lovell and Schmidt, and Bauer.

Work has continued on the "nonparametric" approach, spurred primarily by Färe and his associates, and a group in management science (Banker, Charnes, and Cooper; Charnes et al.) who call the approach data envelopment analysis. A recent survey of new developments is Seiford and Thrall. The two approaches are different in many ways. Banker, Conrad and Strauss; Banker, Charnes, Cooper, and Maindiratta; and Gong and Sickles provide empirical comparisons between the methods. A recent issue of the *Journal of Econometrics* (Lewin and Knox Lovell) is also devoted to both approaches. One disadvantage of the parametric approach is the difficulty of estimating multiproduct production functions. This has led several researchers to propose the use of multiproduct cost structures to evaluate efficiency (Schmidt and Knox Lovell; Kopp and Diewert; Zieschang). While this approach has significant merit, it requires price data and a specific objective function to measure a purely technical phenomenon. An alternative is to estimate the multiproduct distance function (Grosskopf and Hayes) using a composed error term. The nonparametric approach generalizes easily to multiple products and has a straightforward interpretation in terms of the input correspondence. It also allows wide variety in terms of technological assumptions.

The nonparametric approach also allows determination of the inefficiency of any individual firm in the sample. The stochastic frontier approach computes the frontier and an average level of inefficiency. Jondrow et al. have proposed a way to compute the expected value of the technical-inefficiency component of the error given the total value, while Battese and Coelli obtain a best predictor of the inefficiency component for panel data sets. For determination of individual firm efficiency, the nonparametric approach has the upper hand at the moment.

The nonparametric approach does not allow for measurement error or outliers in any meaningful sense. While the work of Varian (1985), and Epstein and Yatchew offers some hope, this is still a major shortcoming of the approach. The stochastic frontier approach is very flexible in this regard and allows an approximate frontier that best fits the data.

The other disadvantage of the nonparametric approach is the lack of any statistical inference. Given the types of measurement and modeler error inherent in economics, this is a flaw that needs to be rectified. Traditional nonparametric statistical methods may provide a way to solve some of these problems. For single-output firms, kernel or semiparametric estimators may provide a way to estimate stochastic frontiers without imposition of functional form. Nonparametric tests could also be developed to consider the results of programming or optimization models used in constructing efficient sets.

There have been numerous applications of the nonparametric approach to efficiency. The early paper by Timmer considered U.S. agriculture using state data as individual observations. Hall and Leveen used the Farrell approach to consider large and small farms in California. Byrnes, Färe, and Grosskopf analyze the efficiency of Illinois strip mines and relate it to optimal scale. Färe, Grabowski, and Grosskopf analyze aggregate Philippine agriculture. Rangan, Grabowski, Aly, and Pasurka apply the technique to U.S. banks, while Grabowski and Pasurka study U.S. farms in 1860. Byrnes, Färe, Grosskopf, and Kraft analyze the relative performance of Illinois grain farms and decompose the efficiency measures to determine that size is a major explanation for differences in efficiency. Färe, Grosskopf, and Kokkelenberg consider plant capacity and technical change in a study of Illinois electric utilities.

Testing for Functional Structure

Tests for functional structure have become popular since the advent of flexible functional forms (Blackorby, Primont, and Russell). Flexible functional forms were an attempt to avoid the maintained hypotheses inherent in traditional forms, such as the Cobb-Douglas (Diewert 1973b; Lau). While these forms have developed immense popularity, they still are not completely general, as exemplified by their separability-inflexibility (Blackorby, Primont, and Russell). Various approaches have been proposed to solve this problem (Pope and Hallam). Nonparametric tests provide one plausible alternative. The techniques of Varian, and Diewert and Parkin allow a nonstatistical test of a variety of functional hypotheses. Chavas and Cox (1988) use aggregate U.S. data and test hypotheses regarding separability and technical change. Fawson and Shumway test separability and jointness hypotheses for ten U.S. agricultural production regions. Barnhart and Whitney provide some comparisons between approaches.

Nonparametric tests on the Varian type are an alternative to traditional parametric tests of functional structure. They have several shortcomings, however. The most critical is the lack of a statistical basis. A comment by Hanoch and Rothschild is useful in this regard:

> We suggest that researchers run these tests on data before they use the data to estimate specific cost or production functions. Doing so will help identify outlying observations and gross inconsistencies, and also provide more insight into the potential usefulness of the data.

The greatest value of the methods may be in performing preliminary data analysis and not in any final hypothesis testing. A second criticism is the inability to deal with outliers or "strange" observations. Points that violate the maintained hypotheses may be isolated cases, but there is no way to weight their significance. The third problem is the ease of use of the tests. Given the simplicity of the procedures, there is a danger that significant research resources will be devoted to applying them indiscriminately to a wide range of possibly interesting, but not terribly important, problems.

Changes in Taste and Technology

A recent application of the nonparametric approach is the changes in taste or technology. As discussed by Varian and others, the reason that a given set of time-series data may not satisfy the axioms of utility maximization is that tastes may change over time. Chalfant and Alston have utilized this point to construct a test for taste change. They analyze data on meat demand and conclude that since it satisfies the revealed-preference axioms, taste change did not occur. They further analyze this issue in a later paper comparing parametric and nonparametric procedures (Alston and Chalfant). Browning has used a similar approach to test aggregate data for consistency with the rational expectations hypothesis. The general rise in income, however, may bias such tests. Specifically, as budget lines shift out over time as income rises, there will be few bundles not consumed at current income that were affordable at previous income levels. Thus, the power of these tests may be low. Chalfant and Alston suggest that income elasticities could be used to investigate the power of this test. This idea was pursued by Sakong, who constructed a compensated test of taste change that tested for consistency of the income-compensated demand curves. In this work, violations of the Afriat axioms were attributed to taste change, only after allowing for changes in income. In effect, his algorithm computes the income elasticities necessary for the data to satisfy the axioms along compensated demand curves.

Chavas and Cox (1990) have used a similar technique to study technical change in the U.S. and Japan. They postulate a particular form of input augmenting technical progress and then test for it nonparametrically. Their test is based on the idea that if the data do not satisfy the postulates of a regular technology and cost minimization, it must be due to technical change. The augmentation coefficients allow for this modification of the technology in order to satisfy the axioms. Thus, the procedure allows for a shifter (technical change) to satisfy the postulates.

Both of these nonparametric approaches are subject to a number of criticisms. The most important is that both explicitly or implicitly attribute all violations of the behavioral and technological postulates to a particular factor, taste or technical change. Given the amount of measurement error in the data and the models, and the presence of other factors, this seems undesirable. This is perhaps even more serious than in the efficiency approach, where no explicit reason is usually given for deviations from the frontier. This approach also suffers from the criticisms of no statistical properties and outliers discussed earlier.

Applications to Finance

A promising area of application of nonparametrics is finance. Varian (1983b) suggested this application long ago. It may be particularly fruitful given the large data series available in finance. It may also be a way to supplement traditional stochastic dominance analysis. Sengupta has provided a way to estimate the portfolio efficiency frontier using this method. It may also be useful in analyzing models with risk. For example, one could determine the level of risk aversion consistent with a particular set of choices. Another application might be to measure the risk premium in financial markets. While this literature is just starting, it may be an important application in the future.

Conclusions

This paper has discussed a variety of nonparametric approaches to economic analysis. A few summary points may be useful.

1. Traditional nonparametric methods, such as order statistics, are underutilized in much of economic research and could be fruitfully exploited in the future, particularly for preliminary data analysis and in analyzing the results of economic models.
2. Series and kernel estimators provide an important way to analyze large data sets. They avoid unnecessary functional specifications but may be of limited use for many data sets and time periods.
3. Semiparametric estimators are a compromise that show great promise. They will require improved computing skills on the part of researchers. They will also require more emphasis on the statistical basis of estimates. Most researchers tend to be spoiled by the assumptions and strong properties of least squares, and have forgotten most of the probability and inference they ever learned.
4. The nonparametric procedures catalogued under the terms data envelopment analysis and re-

vealed preference suffer from a lack of statistical basis. They are nevertheless extremely powerful ways to analyze data in a general framework. They may be preferred to stochastic methods using restrictive functional forms. The search for a statistical basis for this work is a worthwhile one that could have enormous payoff.

5. There is some danger in attributing violations of axioms or hypotheses to any particular factor when several are possible. This is not just a criticism of the nonparametric approach, but of model building in general.

6. There is some danger in the indiscriminate use of nonparametric techniques. It would be unfortunate if they became the latest ''new technique'' in search of a problem.

Perhaps the greatest advantage of nonparametric techniques is that they are data-centered. There is a tendency among economists to be obsessed with models and model building, and to be less concerned with the nature of the underlying data. Nonparametric techniques encourage the active exploration of the data and its central tendencies. To the extent that the use of this approach encourages more attention to data, economics will be a richer science.

References

Afriat, S.N. ''The Construction of Utility Functions from Expenditure Data.'' *International Economic Review* 8(1967): 67–77.

______. ''Efficiency Estimation of Production Functions.'' *International Economic Review* 13(1972):568–98.

______. ''On A System of Inequalities in Demand Analysis: An Extension of the Classical Method.'' *International Economic Review* 14(1973):460–72.

Aigner, D., and S.F. Chu. ''On Estimating the Industry Production Function.'' *American Economic Review* 58(1968): 826–39.

Aigner, D., C.A. Knox Lovell, and P. Schmidt. ''Formulation and Estimation of Stochastic Frontier Production Function Models.'' *Journal of Econometrics* 6(1977):21–37.

Aizcorbe, A.M. ''A Lower Bound for the Power of Nonparametric Tests.'' *Journal of Business and Economic Statistics* 9(1991):463–67.

Alston, J.M., and J.A. Chalfant. ''Can We Take the Con Out of Meat Demand Studies.'' *Western Journal of Agricultural Economics* 16(1991):36–48.

Andrews, D.W.K. ''Asymptotic Normality of Series Estimators for Nonparametric and Semiparametric Regression Models.'' *Econometrica* 59(1991):307–46.

Banker, R.D., A. Charnes, and W.W. Cooper. ''Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis.'' *Management Science* 30(1984): 1078–92.

Banker, R.D., A. Charnes, W.W. Cooper, and A. Maindiratta. ''A Comparison of DEA and Translog Estimates of Production Frontiers Using Simulated Observations from a Known Technology.'' In *Applications of Modern Production Theory: Efficiency and Productivity*, edited by A. Dogramaci and R. Färe. Boston: Kluwer Academic Publishers, 1988.

Banker, R.D., and A. Maindiratta. ''Nonparametric Analysis of Technical and Allocative Efficiencies in Production.'' *Econometrica* 56(1988):1315–32.

Banker, R.D., R. Conrad, and R.P. Strauss. ''A Comparative Application of Data Envelopment Analysis and Translog Methods: An Illustrative Study of Hospital Production.'' *Management Science* 32(1986):30–43.

Barnett, W.A., and P. Yue. ''Semiparametric Estimation of the Asymptotically Ideal Model: The AIM Demand System.'' In *Advances in Econometrics*, vol. 7, edited by G.F. Rhodes and T.B. Fomby. Greenwich: JAI Press, 1988.

Barnett, W.A., and Y.W. Lee. ''The Global Properties of the Minflex Laurent, Generalized Leontief, and Translog Flexible Functional Forms.'' *Econometrica* 53(1985): 1421–37.

Barnhart, S.W., and G.A. Whitney. ''Nonparametric Analysis in Parametric Estimation: An Application to Translog Demand Systems.'' *Review of Economics and Statistics* 70(1988):149–53.

Battese, G.E., and T.J. Coelli. ''Prediction of Firm-level Technical Efficiencies with a Generalized Frontier Production Function and Panel Data.'' *Journal of Econometrics* 38(1988):387–99.

Bauer, P.W. ''Recent Developments in the Econometric Estimation of Frontiers.'' *Journal of Econometrics* 46(1990): 39–56.

Bierens, H. ''Kernel Estimation of Regression Functions.'' In *Advances in Econometrics, Fifth World Congress*, vol. 1, edited by T. Bewley. Cambridge: Cambridge University Press, 1987.

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

Boles, J. ''Efficiency Squared-Efficient Computation of Efficiency Indexes.'' In *Proceedings Thirty-ninth Annual Meeting, Western Farm Economics Association*. 1966.

______. *The 1130 Farrell Efficiency System—Multiple Products, Multiple Factors*. Berkeley: Giannini Foundation of Agricultural Economics, 1971.

Bressler, R.G. ''The Measurement of Productive Efficiency.'' In *Proceedings Thirty-ninth Annual Meeting, Western Farm Economics Association*. 1966.

Bronars, S.G. ''The Power of Nonparametric Tests of Preference Maximization.'' *Econometrica* 55(1987):693–98.

Browning, M. ''A Nonparametric Test of the Life-Cycle Rational Expectations Hypothesis.'' *International Economic Review* 30(1989):979–92.

Byrnes, P., R. Färe, and S. Grosskopf. ''Measuring Productive Efficiency: An Application to Illinois Strip Mines.'' *Management Science* 30(1984):671–81.

Byrnes, P., R. Färe, S. Grosskopf, and S. Kraft. ''Technical Efficiency and Size: The Case of Illinois Grain Farms.'' *European Review of Agricultural Economics* 14, no. 4(1987):367–82.

Campbell, B., and J. Dufour. "Over-rejections in Rational Expectations Models." *Economics Letters* 35(1991):285–90.

Chalfant, J.A. "A Globally Flexible, Almost Ideal Demand System." *Journal of Business and Economic Statistics* 5(1987):233–42.

Chalfant, J.A., and J.M. Alston. "Accounting for Changes in Tastes." *Journal of Political Economy* 96(1988):391–410.

Chamberlain, G. "Asymptotic Efficiency in Estimation with Conditional Moment Restrictions." *Journal of Econometrics* 34(1987):305–34.

______. "Efficiency Bounds for Semiparametric Regression." *Econometrica* 60(1992):567–96.

Charnes, A., W.W. Cooper, B. Golany, L. Seiford, and J. Stutz. "Foundations of Data Envelopment Analysis for Pareto-Koppmans Efficient Empirical Production Functions." *Journal of Econometrics* 30(1985):91–107.

Chavas, J., and T.L. Cox. "A Nonparametric Analysis of Agricultural Technology." *American Journal of Agricultural Economics* 70(1988):303–10.

______. "A Nonparametric Analysis of Productivity: The Case of U.S. and Japanese Manufacturing." *American Economic Review* 80(1990):450–64.

Coslett, S.R. "Efficiency Bounds For Distribution Free Estimators of the Binary Choice Model and the Censored Regression Models." *Econometrica* 55(1987):559–85.

Deaton, A.S. "Rice Prices and Income Distribution in Thailand." *Economic Journal* 99(1989):1–37.

Diewert, W.E. "Afriat and Revealed Preference Theory." *Review of Economic Studies* 40(1973a):419–26.

______. "Functional Forms for Profit and Transformation Functions." *Journal of Economic Theory* 6(1973b):284–316.

______. "Hick's Aggregation Theorem and the Existence of a Real Value Added Function." In *Production Economics: A Dual Approach to Theory and Practice,* edited by M. Fuss and D. McFadden. Amsterdam: North Holland, 1978.

Diewert, W.E., and C. Parkin. "Linear Programming Tests of Regularity Conditions for Production Functions." In *Quantitative Studies on Production and Prices,* edited by W. Eichorn, R. Henn, K. Newmann, and R.W. Shephard. Wurzburg-Wien: Physica-Verlag, 1983.

______. "Tests for the Consistency of Consumer Data." *Journal of Econometrics* 30(1985):127–47.

Diewert, W.E., and T.J. Wales. "Flexible Functional Forms and Global Curvature Conditions." *Econometrica* 55(1987):43–68.

Engle, R.F., C.W.J. Granger, J. Rice, and A. Weiss. "Semiparametric Estimates of the Relation Between Weather and Electricity Sales." *Journal of the American Statistical Association* 81(1986):310–20.

Epanechinikov, V.A. "Nonparametric Estimation of a Multivariate Probability Density." *Theory of Probability and Its Applications* 14(1969):153–58.

Epstein, L.G., and A.J. Yatchew. "Nonparametric Hypothesis Testing Procedures and Applications to Demand Analysis." *Journal of Econometrics* 30(1985):149–69.

Färe, R. *Fundamentals of Production Theory*. New York: Springer-Verlag, 1988.

Färe, R., R. Grabowski, and S. Grosskopf. "Technical Efficiency of Philippine Agriculture." *Applied Economics* 17(1985):205–14.

Färe, R., S. Grosskopf, and C.A. Knox Lovell. *The Measurement of Efficiency of Production*. Boston: Kluwer-Nijhoff, 1985.

Färe, R., S. Grosskopf, and E.C. Kokkelenberg. "Measuring Plant Capacity, Utilization and Technical Change: A Nonparametric Approach." *International Economic Review* 30(1989):655–66.

Farrell, M.J. "The Measurement of Economic Efficiency." *Journal of the Royal Statistical Society* 120, Pt. III(1957): 252–81.

Fawson, C., and C.R. Shumway. "A Nonparametric Investigation of Agricultural Production Behavior for U.S. Subregions." *American Journal of Agricultural Economics* 70(1988):311–17.

Friedman, M. "The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance." *Journal of the American Statistical Association* 32(1937):675–701.

Gallant, A.R. "On the Bias in Flexible Functional Forms and an Essentially Unbiased Form: The Fourier Flexible Form." *Journal of Econometrics* 15(1981):211–45.

______. "Unbiased Determination of Production Technologies." *Journal of Econometrics* 20(1982):285–323.

______. "The Fourier Flexible Form." *American Journal of Agricultural Economics* 66(1984):204–8.

Gallant, A.R., and D.W. Nychka. "Semi-nonparametric Maximum Likelihood Estimation." *Econometrica* 55(1987): 363–90.

Gallant, A.R., D.A. Hsieh, and G.E. Tauchen. "On Fitting a Recalcitrant Series: The Pound/Dollar Exchange Rate, 1974–1983." In *Nonparametric and Semiparametric Methods in Economics and Statistics,* edited by W.A. Barnett, J. Powell, and G. Tauchen. Cambridge: Cambridge University Press, 1991.

Gibbons, J.D. *Nonparametric Statistical Inference*. New York: Marcel Dekker, 1985.

Gong, B., and R. Sickles. "Finite Sample Evidence on the Performance of Stochastic Frontiers and Data Envelopment Analysis Using Panel Data." *Journal of Econometrics* 51(1992):259–84.

Grabowski, R., and C. Pasurka. "Farmer Education and Economic Efficiency: Northern Farms in 1860." *Economics Letters* 28(1988):315–20.

Grenander, U. *Abstract Inference*. New York: Wiley, 1981.

Grosskopf, S., and K. Hayes. "Local Public Sector Bureaucrats and Their Input Choices." *Journal of Urban Economics* (forthcoming).

Hall, B.F., and E.P. Leveen. "Farm Size and Economic Efficiency: The Case of California." *American Journal of Agricultural Economics* 60(1978):589–600.

Han, A.K. "A Nonparametric Analysis of Transformations." *Journal of Econometrics* 35(1987a):191–209.

______. "Nonparametric Analysis of a Generalized Regression Model." *Journal of Econometrics* 35(1987b):303–16.

Hanoch, G., and M. Rothschild. "Testing the Assumptions of Production Theory." *Journal of Political Economy* 80(1972):256–74.

Hardle, W. *Applied Nonparametric Regression*. Cambridge: Cambridge University Press, 1990.

Hardle, W., W. Hildenbrand, and M. Jerison. "Empirical Evidence on the Law of Demand." *Econometrica* 59(1991): 1525–50.

Holmes, J.M., and P.A. Hutton. "On the Causal Relationship Between Government Expenditures and National Income." *Review of Economic Statistics* 72(1990):87–95.

Holt, M.T., and G. Moschini. "Alternative Measures of Risk in Commodity Supply Models: An Analysis of Sow Farrowing Decisions in the United States." *Journal of Agricultural and Resource Economics* 17(1992):1–12.

Hong, Y., and A. Pagan. "Some Simulation Studies of Nonparametric Estimators." *Empirical Economics* 13(1988): 251–66.

Horowitz, J.L. "A Smoothed Maximum Score Estimator for the Binary Response Model." *Econometrica* 60(1992): 505–32.

Hotelling, H., and M.R. Pabst. "Rank Correlation and Tests of Significance Involving No Assumptions of Normality." *Annuals of Mathematical Statistics* 7(1936):29–43.

Jondrow, J., C.A. Knox Lovell, I.S. Materov, and P. Schmidt. "On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model." *Journal of Econometrics* 19(1982):233–38.

Kendall, M. "A New Measure of Rank Correlation." *Biometrika* 30(1938):81–93.

______. *Rank Correlation Methods*. London: Griffin, 1948.

Knox Lovell, C.A., and P. Schmidt. "A Comparison of Alternative Approaches to the Measurement of Productive Efficiency." In *Applications of Modern Production Theory: Efficiency and Productivity*, edited by A. Dogramaci and R. Färe. Boston: Kluwer Academic Publishers, 1988.

Kopp, R.J., and W.E. Diewert. "The Decomposition of Frontier Cost Function Deviations into Measures of Technical and Allocative Efficiency." *Journal of Econometrics* 19(1982):319–31.

Lau, L.J. "Applications of Duality Theory: A Comment." In *Frontiers of Quantitative Economics II*, edited by M. Intriligator and D. Kendrick. Amsterdam: North Holland, 1974.

Lee, L. "Semiparametric Nonlinear Least Squares Estimation of Truncated Regression Models." *Econometric Theory* 8(1992):52–94.

Lehmann, E.L. *Nonparametrics: Statistical Methods Based on Ranks*. San Francisco: Holden-Day, 1975.

Lewin, A.Y., and C.A. Knox Lovell, eds. "Frontier Analysis: Parametric and Nonparametric Approaches." *Journal of Econometrics* 46(1990).

Maindiratta, A. "Largest Size-Efficient Scale and Size Efficiencies of Decision-making Units in Data Envelopment Analysis." *Journal of Econometrics* 46(1990):57–72.

Mann, H.B., and D.R. Whitney. "On a Test of Whether One of Two Random Variables is Stochastically Larger than the Other." *Annuals of Mathematical Statistics* 18(1947): 50–60.

Manski, C.F. "Adaptive Estimation of Nonlinear Regression Models." *Econometric Reviews* 3(1984):145–94.

______. "Semiparametric Analysis of Discrete Response: Asymptotic Properties of the Maximum Score Estimator." *Journal of Econometrics* 27(1985):313–34.

______. "Semiparametric Analysis of Random Effects Linear Models from Binary Panel Data." *Econometrica* 55(1987):357–62.

Matzkin, R.L. "A Nonparametric Rank Correlation Estimator." In *Nonparametric and Semiparametric Methods in Economics and Statistics*, edited by W.A. Barnett, J. Powell, and G. Tauchen. Cambridge: Cambridge University Press, 1991a.

______. "Semiparametric Estimation of Monotone and Concave Utility Functions for Polychotomous Choice Models." *Econometrica* 59(1991b):1315–28.

McCurdy, T.H., and T. Stengos. "A Comparison of Risk-Premium Forecasts Implied by Parametric versus Nonparametric Conditional Mean Estimators." *Journal of Econometrics* 52(1992):225–44.

Meeusen, W., and J. van den Broeck. "Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error." *International Economic Review* 18(1977):435–44.

Moschini, G. "Nonparametric and Semiparametric Estimation: An Analysis of Multiproduct Returns to Scale." *American Journal of Agricultural Economics* 72(1990):589–96.

______. "Testing for Preference Change in Consumer Demand: An Indirectly Separable Semiparametric Model." *Journal of Business and Economic Statistics* 9(1991):111–17.

Nadaraya, E.A. "On Estimating Regression." *Theory of Probability Applications* 9(1964a):141–42.

______. "Some New Estimates for Distribution Functions." *Theory of Probability Applications* 9(1964b):497–500.

Newey, W.K. "Adaptive Estimation of Regression Models via Moment Restrictions." *Journal of Econometrics* 38(1988):301–9.

______. "Semiparametric Efficiency Bounds." *Journal of Applied Econometrics* 5(1990):99–135.

Pagan, A.R., and A. Ullah. "The Econometric Analysis of Models with Risk Terms." *Journal of Applied Econometrics* 3(1988):87–105.

Pagan, A.R., and G.W. Schwert. "Testing for Covariance Stationarity in Stock Market Data." *Economics Letters* 33(1990):165–70.

Pagan, A.R., and M.R. Wickens. "A Survey of Some Recent Econometric Methods." *Economic Journal* 99(1989): 962–1025.

Pagan, A.R., and Y.S. Hong. "Nonparametric Estimation and the Risk Premium." In *Nonparametric and Semiparametric Methods in Economics and Statistics*, edited by W.A. Barnett, J. Powell, and G. Tauchen. Cambridge: Cambridge University Press, 1991.

Pope, R.D. "Estimating Functional Forms with Special Reference to Agriculture: Comments." *American Journal of Agricultural Economics* 66(1984):223–24.

Pope, R.D., and A. Hallam. "Separability Testing in Production Economics." *American Journal of Agricultural Economics* 70(1988):142–52.

Prakasa Rao, B.L.S. *Nonparametric Functional Estimation*. Orlando: Academic Press, 1983.

Randles, R.H., and D.A. Wolfe. *Introduction to the Theory of Nonparametric Statistics*. New York: Wiley, 1979.

Rangan, N., R. Grabowski, H.Y. Aly, and C. Pasurka. "The Technical Efficiency of U.S. Banks." *Economics Letters* 28(1988):169–75.

Rilstone, P. "Nonparametric Hypothesis Testing with Parametric Rates of Convergence." *International Economic Review* 32(1991):209–27.

Robinson, P.M. "Nonparametric Estimators for Time Series." *Journal of Time Series Analysis* 4(1983):185–207.

______. "Semiparametric Econometrics: A Survey." *Journal of Applied Econometrics* 3(1988):35–51.

______. "Consistent Nonparametric Entropy-Based Testing." *Review of Economic Studies* 58(1991):437–53.

Rosenblatt, M. "Remarks on Some Nonparametric Estimates of a Density Function." *Annuals of Mathematical Statistics* 27(1956):832–7.

Russell, R.R. "Measures of Technical Efficiency." *Journal of Economic Theory* 35(1985):109–26.

Sakong, Y. "Essays in Nonparametric Measures of Changes in Taste and Hedging Behavior with Options." Ph.D. diss., Iowa State University, 1991.

Sakong, Y., and D.J. Hayes. "A Test for the Consistency of Demand Data with Revealed Preference Theory." *American Journal of Agricultural Economics* (forthcoming).

Schmidt, P. "Frontier Production Functions." *Econometric Reviews* 4(1986):289–328.

Schmidt, P., and C.A. Knox Lovell. "Estimating Technical and Allocative Inefficiency Relative to Stochastic Production and Cost Frontiers." *Journal of Econometrics* 9(1979):363–66.

Seiford, L.M., and R.M. Thrall. "Recent Developments in Data Envelopment Analysis: The Mathematical Programming Approach to Frontier Analysis." *Journal of Econometrics* 46(1990):7–38.

Seitz, W.D. "The Measurement of Efficiency Relative to a Frontier Production Function." *American Journal of Agricultural Economics* 52(1970):505–11.

Sengupta, J.K. "Nonparametric Test of the Efficiency of Portfolio Investment." *Zeitschrift für Nationalolonomie* 50(1989):1–15.

Sentana, E., and S. Wadhwani. "Semiparametric Estimation and the Predictability of Stock Market Returns: Some Lessons from Japan." *Review of Economic Studies* 58(1991): 547–63.

Shephard, R.W. *Indirect Production Functions*. Mathematical Systems in Economics, no. 10. Meisenheim Am Glad: Verlag Anton Hain, 1974.

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

Stein, C. "Efficient Nonparametric Testing and Estimation." In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley: University of California Press, 1956.

Stoker, T.M. "Test of Additive Derivative Constraints." *Review of Economic Studies* 56(1989):535–52.

Stone, C.J. "Adaptive Maximum Likelihood Estimation of a Location Parameter." *Annals of Statistics* 3(1975):267–84.

Timmer, C.P. "Using a Probabilistic Frontier Production Function to Measure Technical Efficiency." *Journal of Political Economy* 79(1971):776–94.

Tolley, H.D., and R.D. Pope. "Testing for Stochastic Dominance." *American Journal of Agricultural Economics* 70(1988):693–700.

Tsur, Y. "On Testing for Revealed Preference Conditions." *Economics Letters* 31(1989):359–62.

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

Varian, H.R. "The Nonparametric Approach to Demand Analysis." *Econometrica* 50(1982):945–73.

______. "Nonparametric Tests of Consumer Behavior." *Review of Economic Studies* 50(1983a):99–110.

______. "Nonparametric Tests of Models of Investor Behavior." *Journal of Financial and Quantitative Analysis* 18(1983b):269–78.

______. "The Nonparametric Approach to Production Analysis." *Econometrica* 52(1984):579–97.

______. "Nonparametric Analysis of Optimizing Behavior with Measurement Error." *Journal of Econometrics* 30(1985):445–58.

______. "Goodness of Fit in Optimizing Models." *Journal of Econometrics* 46(1990):125–40.

Vinod, H.D., and A. Ullah. "Flexible Production Function Estimation by Nonparametric Kernel Estimators." In *Advances in Econometrics*, vol. 7, edited by G.F. Rhodes and T.B. Fomby. Greenwich: JAI Press, 1988.

Watson, G.S. "Smooth Regression Analysis." *Sankhya Series A* 26(1964):359–72.

White, H., and J.M. Wooldridge. "Some Results on Sieve Estimation with Dependent Observations." In *Nonparametric and Semiparametric Methods in Economics and Statistics*, edited by W.A. Barnett, J. Powell, and G. Tauchen. Cambridge: Cambridge University Press, 1991.

Wilcoxin, F. "Individual Comparisons by Ranking Methods." *Biometrics* 1(1945):80–83.

Yatchew, A. "Nonparametric Regression Model Tests Based on Least Squares." *Econometric Theory* (forthcoming, 1992).

Zieschang, K.D. "A Note on the Decomposition of Cost Efficiency into Technical and Allocative Components." *Journal of Econometrics* 23(1983):401–5.