Nonparametric Analysis of Production Efficiency: Discussion

Conrado M. Gempesaw II

When I was invited to be a discussant in this session dealing with nonparametric analysis of production and consumption, my initial response was rather mixed. On one hand, I teach a graduate-level course in applied statistics, and one of the topics I cover is nonparametric statistics and its application to the biological sciences. Hence, I thought it would be interesting to relate what I cover in class with what is used in applied agricultural production economics research. On the other hand, I am not familiar with this approach since I have not used the nonparametric programming method for efficiency analysis in my past research. Thus, together with my graduate students, we started applying the nonparametric production efficiency analysis in order to gain a better understanding of this approach. The deeper we got into the research topic, I realized that there was very little relationship between what is called nonparametric statistics by statisticians and nonparametric methods used to analyze production efficiency.

In this discussion paper, I will focus mainly on the papers by Arne Hallam and Hyunok Lee and leave the consumption paper of James Chalfant and Julian Alston to the other discussant, Laura Blanciforti. My discussion will be organized in three sections. The first section presents some comments on the use of nonparametric methods. The second section discusses a technique that I would like to propose as an alternative in efficiency measurement. The third section raises the issue whether the set of methodologies in efficiency measurement is becoming an end in itself.

Nonparametric Efficiency: Some Empirical Comments

A considerable amount of literature has focused on nonparametric production efficiency measurement over the last decade. An excellent overview of the concepts and application of this approach can be found in the paper written by Arne Hallam. He begins with a classification of four types of analysis that are lumped under the heading "nonparametric." Briefly, they are nonparametric inference, nonparametric density estimation, semi-nonparametric estimation, and nonparametric measurement of production efficiency. Hallam gives an outstanding overview on these four types of analysis, and whatever comments I have are basically clarifications to what he has already mentioned.

First, parametric statistical inference is generally based on the assumption that the random samples are selected from some underlying (e.g., normal) distribution. Most of parametric statistical tests are reliable even if there are slight departures from normality, as long as sample sizes are large. On the other hand, nonparametric statistical inference often assumes no knowledge about the distribution of the underlying population. Nonparametric inference has gained popularity because of several reasons. The computations are easy to estimate. The data need not be measured quantitatively but could be in a qualitative format. The data could also be in an ordinal ranking. Nonparametric inference also does not have as many restrictive assumptions as parametric statistical inference. On the other hand, nonparametric inference is deficient in the sense that it does not utilize all the information in the sample and, thus, will be less efficient than parametric inference. Statisticians would indicate that if the mean is the central measure in parametric statistics, the median is considered the central measure in nonparametric statistics. In this manner, it can be shown that nonparametric inference does not utilize all information in the sample.

Second, as noted by Hallam, nonparametric efficiency analysis is called "nonparametric" in the sense that no functional form is assumed and has no direct relationship with nonparametric statistical inference. Although this is generally true, it could be said that nonparametric efficiency analysis may be similar to nonparametric statistical in-

Conrado M. Gempesaw II is an associate professor, Delaware Agricultural Experiment Station, Department of Food and Resource Economics, University of Delaware, College of Agricultural Sciences, Newark, DE.
Published as Miscellaneous Paper no. 1450 of the Delaware Agricultural Experiment Station.

ference in that not all information is used in nonparametric efficiency analysis. This is because in nonparametric programming analysis we are not able to specify a functional relationship between outputs and inputs in a meaningful manner. The only relationship between the outputs and inputs is that outputs are used as the upper-bound constraints, with inputs specified as the lower-bound constraints. Another similarity is that nonparametric statistical inference is distribution-free, which is generally true for nonparametric programming methods.

Third, Hallam points out the disadvantages of nonparametric efficiency analysis in that it does not allow for measurement errors and lacks statistical inference. Another major disadvantage we have found in our own research work is that it is very sensitive to input aggregation assumptions. For example, we estimated the technical efficiency of German dairy farms using the Färe and Grosskopf linear-programming method. A discussion of the data set is available in Fernandez-Cornejo. Initially, we specified seven input categories, then we respecified the model with five input categories, and finally, we reestimated the model with three input categories. As shown in Table 1, as inputs are further disaggregated, the farms achieve higher technical-efficiency levels. Similar findings were also reported by Thomas and Tauer in their efficiency study of New York dairy farms. If this is indeed true, then researchers can improve their efficiency measures simply by disaggregating input categories—not a comforting result, particularly given the amount of literature focused on nonparametric efficiency measurement over the last ten years.

Fourth, Hallam reviews parametric approaches to efficiency analysis and notes the use of the stochastic frontier approach as a flexible technique in measuring the frontier production function, which provides for a meaningful estimate of the measurement error. Although in theory the stochastic frontier approach has a solid foundation, our own research experience with this technique has not been encouraging. In most cases, the estimation technique reveals that most of the errors are symmetric. Although this may not necessarily be true, problems in data measurement can cause such results. Considering that data are always measured with some degree of error, the researcher is left to decide whether producers are truly efficient or if the extent of the data errors has affected the results, thus leading to erroneous efficiency conclusions. Furthermore, there are several instances when the estimation technique does not converge, leading the researcher with questions on what to do next. Of course, it is possible that "cleaning" the data could solve both empirical problems, but that is a rather discouraging alternative.

Fifth, another important empirical observation we found in our research work is that using the stochastic frontier generally results in lower measures of technical efficiency compared to the nonparametric programming method. Although this is understandable since the nonparametric method does not allow for external shocks, it is still not encouraging for practitioners to be using techniques that produce different measures of efficiency for the same body of data. In fact, similar findings were reported by Kalaitzandonakes, Wu, and Ma. The authors compared technical efficiencies for Missouri grain farms using different methods and found that the nonparametric programming method resulted in markedly high measures

Table 1. Technical Efficiency Estimates for German Dairy Farms by Various Input Aggregation

Size	1983			1985			1987			1989		
of Farm[a]	Seven[b]	Five	Three	Seven	Five	Three	Seven	Five	Three	Seven	Five	Three
I	100	99	97	100	99	97	100	100	99	100	100	99
II	99	97	94	99	98	95	99	98	96	99	98	95
III	98	96	92	99	99	95	100	99	97	99	98	96
IV	100	99	97	100	99	98	100	99	98	100	99	98

[a]Size of farm was based on the amount of annual milk output produced. The first classification (I) included farms producing less than 50,000 kilograms of milk; the second classification (II) included farms producing between 50,000 to 100,000 kilograms of milk; the third classification (III) included farms producing 100,000 to 150,000 kilograms of milk; and the fourth classification (IV) included farms producing over 150,000 kilograms of milk.

[b]Input aggregation was classified as follows: seven inputs included feed, crops, hired labor, miscellaneous, capital, land, and family labor; five inputs included feed, a combination of crops and miscellaneous, capital, land, and a combination of hired and family labor; three inputs included feed, a combination of crops, hired labor, and miscellaneous, and a combination of capital, land, and family labor.

of technical efficiency relative to the stochastic frontier measures. The question is, which method is more accurate?

Lee's paper focuses on the application of nonparametric programming methods to analyze the consequence of economic, regulatory, and contractual constraints on firm practices. She discusses these applications particularly with respect to the notion of an expenditure-constrained profit function. What is interesting is that she points out some shortcomings of applying the nonparametric programming technique for most of the proposed applications. For example, she points out that in an expenditure-constrained profit-maximization problem, rejection of the profit-maximization hypothesis could be caused by the presence of allocative inefficiency or by the existence of expenditure constraints. The same shortcoming can also be pointed out for other applications of profit maximization as constrained by rate-of-return regulations, international-trade restrictions, and other contractual constraints.

The paper by Lee presents an excellent, but brief overview of the use of nonparametric programming techniques in a frontier framework to calculate the relative efficiency of firms. She explains that since we measure relative efficiencies, we implicitly presume that firms using the same input mix can produce two different output levels. Lee provides two very interesting reasons why this phenomenon could happen. First, different firms in the same industry have different technologies. Second, different firms in the same industry have the same technology, but some firms use their inputs inefficiently. She argues that the first reason is rather weak because it violates the perfect-information postulate of neoclassical theory. The second reason is also insufficient because it violates the assumption of rationality. Thus, she concludes that outcomes from nonparametric programming methods may be caused by several possible reasons, and their robustness is left open to question. I think that researchers should devise a technique to measure efficiencies by allowing for firms to have different technologies and at the same time measure the input-utilization inefficiency of a firm relative to itself and not based on other firms. This is a topic I would like to briefly discuss in the next section.

An Alternative Technique to Efficiency Measurement

In an article published in this issue of the *Journal,* Lass and Gempesaw propose the use of the random coefficient regression (RCR) method as an alternative technique to economic-efficiency measurement. The reader is referred to the cited reference for the full development of the model. The random coefficient regression approach to efficiency measurement allows the researcher to avoid the two limitations raised by Lee on the use of nonparametric programming to calculate relative-efficiency measures. The RCR approach allows the analyst to estimate separate production technologies for each firm, thereby avoiding having to compare efficiency measures across firms. Furthermore, because parameter estimates are provided for each individual firm, it is possible to compare whether firms in the same industry have different technologies.

Now let us return to the two arguments presented by Lee. If the first argument that different firms in the same industry have different technologies is correct, then efficiency measures from previous approaches are due to differences in technology. This means that what is being measured are differences in technologies and not inefficiencies. What about the perfect-information postulate of neoclassical economic theory? It is obvious that markets may approximate competitive conditions, but the ideal case may not exist. In fact, most of the recent empirical research in industrial organization is focused towards models of market imperfection. In addition, using the RCR approach allows the researcher to test for the second limitation raised by Lee; that is, firms may have the same technology, but some may use their inputs inefficiently. Since optimal input demand can be derived using the first-order condition for profit maximization, one can then test whether firms are economically inefficient by equating the varying input coefficient with the expenditure-to-revenue ratio. This is similar to the marginal value product equated to the input price rule. One advantage of this approach is that we are able to determine the efficiency of each input utilized, and firms that tend to over- or underutilize certain inputs can be identified.

Is Methodology Becoming an End in Itself?

Although the question of whether methodology is becoming an end in itself can be raised not only in the use of different methods in efficiency measurement, I would like to concentrate my comments on the topic being discussed. Almost 30 years ago, Don Paarlberg published an article in the *Journal of Farm Economics* asking agricultural economists whether methodology is becoming an end in itself. He clearly described the problem by stating that ". . . the criterion for the choice of method should be its usefulness in solving the problem, not the degree of its sophistication, nor the date it was first used . . . the method should not become an end in itself" (p. 1388). Furthermore, he states that "we delude ourselves into thinking that the new methods, being mathematical, shelter us from the hazards of human judgment, not realizing that the choice of the model is a piece of subjective judgment and predetermines the outcome of the research" (p. 1390).

What Paarlberg discussed 30 years ago is still very relevant today, when researchers have to decide which method to use in measuring firm efficiency. Although studies have shown that the use of the nonparametric programming technique tends to produce efficiency measures that are higher than those of other parametric approaches, the proliferation in the use of this method continues in the literature. Furthermore, the results have been shown to be very sensitive to input-aggregation assumptions.

Although I do not have a prescription as to how this problem might be solved other than to offer another methodology, I would like to close this discussion by citing a study on technical efficiency conducted three decades ago. In an article published in the *Journal of Farm Economics*, Kadlec and House estimated technical efficiency by taking actual output and dividing it by predicted output from a regression equation of output against input. The index derived reflects the percentage that actual output is of predicted output and was labeled as a technical-efficiency index. The authors argue that although "this measure may be criticized because of input measurement error, and regression fit may bias the index, and while a more independent measure may be more desirable, it is probably closely associated with technical efficiency" (p. 1430). Although we might think this approach is relatively ad hoc and erroneous, I am not sure whether the methodologies we have at present in measuring efficiency are any better, considering that we obtain different efficiency measures based on different methodologies.

References

Färe, R., and S. Grosskopf. "A Nonparametric Cost Approach to Scale Efficiency." *The Scandinavian Journal of Economics* 87(1985):594–604.

Fernandez-Cornejo, J. "Dynamic Structure of Multiproduct Firms: An Application to German Agriculture." Ph.D. diss., University of Delaware, 1990.

Hallam, Arne. "A Brief Overview of Nonparametric Methods in Economics." Invited paper presented at the joint American Agricultural Economics Association–Northeast Agricultural and Resource Economics Association annual meeting, Baltimore, MD, August 1992.

Kadlec, J., and A. House. "The Effect of Technical Efficiency on Optimum Size." *Journal of Farm Economics* 44(1962):428–32.

Kalaitzandonakes, N., S. Wu, and J. Ma. "The Relationship Between Technical Efficiency and Firm Size Revisited." Selected paper presented at the Southern Agricultural Economics Association annual meeting, Lexington, KY, February 1992.

Lass, D., and C. Gempesaw. "Estimation of Firm-Varying, Input-Specific Efficiencies in Dairy Production." *Northeastern Journal of Agricultural and Resource Economics* 21(1992):142–150.

Lee, H. "Recent Applications of Nonparametric Programming Methods." Invited paper presented at the joint American Agricultural Economics Association–Northeast Agricultural and Resource Economics Association annual meeting, Baltimore, MD, August 1992.

Paarlberg, D. "Is Methodology Becoming An End In Itself." *Journal of Farm Economics* 45(1963):1386–92.

Thomas, A., and L. Tauer. "The Causes of Economic Inefficiencies in New York Dairy Farms." Agricultural Economics Staff Paper no. 91-16. Department of Agricultural Economics, Cornell University, June 1991.