Nonparametric Estimation of Production Functions Subject to Cost-Minimization and Profit-Maximization

David Dole
Agricultural and Resource Economics
The University of Western Australia
Nedlands, Western AUSTRALIA 6907
ddole@cyllene.uwa.edu.au
www.uwa.edu.au/cyllene/ddole

Abstract
Nonparametric methods for testing economic hypotheses are extremely powerful, able to detect even the slightest departures from the given hypothesis. This power becomes a disadvantage when such departures are due merely to natural random variation in the data. This paper considers the application of these methods to production data with output subject to random variation. The paper focuses specifically on production under cost-minimization and profit-maximization. Nonparametric theory of production under these conditions dictates constraints on output and costs. When output is subject to random variation, these constraints can be used to estimate the output that is consistent with the given form of production, resulting in a nonparametric estimate of the production function. Monte Carlo tests indicate that, when the behavioural model is true, nonparametric estimates are competitive with estimates from parametric models, even when the parametric family includes the true production function.

The complete version of this paper is available at the web address above.
I. Introduction

Almost any hypothesis about microeconomic behavior can, in principle, be investigated with a nonparametric method. For example, nonparametric methods can determine whether data on a production process are consistent with cost-minimization, or profit-maximization, or constant returns to scale, and a variety of other common economic hypotheses. (Afriat, 1972, Hanoch and Rothschild, 1972, and Varian, 1984) These methods are as powerful as they are widely applicable. Parametric methods are often unsatisfying, or even unconvincing, since inferences are always contingent on the validity of the parametric model adopted in the analysis: a parametric test of an hypothesis can be rejected if either the hypothesis or the model is incorrect. But nonparametric methods are completely free from such limitations: if the hypothesis is rejected, the only valid inference is that the data are inconsistent with the hypothesis.

One of the most remarkable things about nonparametric methods is that they are rarely used. Perhaps one reason for this, paradoxically, lies in the great advantage of these methods -- that they are without limitation, and are unforgiving even of natural random variation in the data. For example, a nonparametric test of cost-minimization can reject the hypothesis even when it is true, if output is subject to random variation. Parametric methods, on the other hand, have been tailored specifically to account for different forms of random variation. Although the use of nonparametric methods is gradually increasing, the continuing popularity of parametric methods reveals that economists prefer the specific treatment of random variation in parametric methods to the freedom from limitation in nonparametric ones.

In fact nonparametric methods are quite amenable to incorporating random variation, despite both appearances to the contrary and the lack of specific treatments in the literature. Hanoch and Rothschild (1972) originally presented their methods in the context of data observed without error, but they also noted that the methods can be extended to handle data subject to random error. Varian (1984) developed and extended the work of Afriat (1972) and Hanoch and Rothschild (1972) in the context of data observed without error, but also went on to consider nonparametric production analysis with errors in inputs (Varian, 1985) and goodness-of-fit (Varian, 1990).

This paper builds on this previous work by considering the application of nonparametric methods to production data when output is subject to random variation, and when both prices and inputs are observed without error -- the usual specification when estimating a production
involving both efficient and inefficient firms, the ranking process was able to identify only a small proportion of the inefficient firms. The methods discussed here should be complemented with other means of diagnosing inefficiency, but that is beyond the scope of this paper. Regressing output on the efficiency ranks should work, in principle, only if production is indeed efficient; when inefficient firms are present in the data, a standard regression of output on inputs would seem to be preferable. However, simulations indicate that regression on the efficiency ranks remains effective even if all inefficient firms are not identified.

Section III considers production under profit-maximization (hereafter called "optimal production"). This form of production gives rise to upper and lower bounds on the differences between the output of all pairs of firms, amounting to \(N^2-N\) inequality restrictions from \(N\) observations. Computation under this many restrictions is infeasible except when \(N\) is small. However, Monte Carlo tests indicate that excellent estimates can be obtained by filtering the data through the restrictions. The resulting estimates are not guaranteed to be completely consistent with profit-maximization, but the simulations indicate that there is very little improvement to be obtained from a comprehensive estimation procedure.

In fact, the restrictions implied by profit-maximization are so strong that in most of the simulations, the filtered output has a lower mean-squared approximation error than a regression on the true production function. In other words, knowing that firms maximize profits can amount to more knowledge about production than if the production function was known up to a linear transformation. This is obviously unreasonable, and so emphasizes that absolute optimality is very unlikely to hold in practice (as is already well known).

The methods presented here are especially applicable to data from agricultural surveys, since agricultural output is typically exposed to uncontrolled environmental effects, and since detailed information can be obtained on inputs. Section III presents an illustration of the methods to data on soybean production in Thailand. The data requirements of the methods are the same as that for the dual approach, and the former can complement the latter. Errors in output can complicate the use of cost functions, since output is a dependent variable. Regression on the efficiency ranks can provide an estimate of output that is consistent with cost-minimization, and so the cost function can be estimated without this errors-in-variables problem. The paper concludes in section IV with a discussion of implications and areas for further research.