Stochastic frontier models

Wang, Hung-Jen

18 March 2006

Stochastic Frontier Models

Prepared for The New Palgrave Dictionary of Economics, 2nd Edition

Hung-Jen Wang*

This Version: March 18, 2006

The stochastic frontier model was first proposed by Aigner et al. (1977) and Meeusen and van den Broeck (1977) in the context of production function estimation. The model extends the classical production function estimation by allowing for the presence of technical inefficiency. The idea is that although the production technology is common knowledge to a group of producers, the efficiency in using that technology in the production process may vary by producers, with the degree of efficiency depending possibly on factors such as experience, management skills, etc.. Given the technology, a fully efficient producer(s) may realize the full potential of the technology and obtain the maximum possible output for given inputs, while less efficient producers see their output fall short of the maximum possible level. Therefore, the underlying technology defines a *frontier* of production, and actual outputs observed in the data fall below the frontier because of the presence of technical inefficiency.

A stochastic production frontier model can be specified as

$$\ln y_i = \ln y_i^* - u_i, \quad u_i \geq 0, \tag{1}$$

$$\ln y_i^* = f(\boldsymbol{x}_i; \boldsymbol{\beta}) + v_i, \tag{2}$$

where y_i is the observed output of producer i, y_i^* is the potential output which is subject to a zero-mean random error v_i, $\boldsymbol{x}_i$ and $\boldsymbol{\beta}$ are vectors of inputs and the corresponding coefficients, respectively, and $u_i \geq 0$ is the effect of technical inefficiency. Equation (2) defines the *stochastic frontier* of the production function; it is stochastic because of v_i. Given that $u_i \geq 0$, observed log of output ($\ln y_i$) is bounded below the frontier. The value of $100 * u_i$ is the percentage by which output can be increased using the same inputs if production is fully efficient. The model without u_i reduces to the classical specification of a production function.

*Institute of Economics, Academia Sinica, Taipei 115, Taiwan. Tel: +886-2-2782-2791 ext.323; fax: +886-2-2735-6006. *E-mail address:* hjwang@econ.sinica.edu.tw.

A popular empirical strategy in estimating the above model is to impose distributional assumptions on u_i and v_i, from which a likelihood function can be derived and estimated. For instance, one may assume that

$$v_i \sim N(0,\ \sigma^2), \tag{3}$$

$$u_i \sim N^+(\mu, \sigma_u^2), \tag{4}$$

where $N^+(\cdot)$ indicates the positive truncation of a normal distribution. The positive truncation gives non-negative values of u_i and hence ensures that firms are constrained by the technology frontier. By making μ and/or σ_u^2 functions of observables (such as ages and years of schooling), one can model the determinants of inefficiency.

The distribution assumption of (4) encompasses many of the models in the literature as special cases. For instance, the half-normal distribution of u_i proposed by Aigner et al. (1977) is obtained by restricting $\mu = 0$ and σ_u^2 to be a constant. The half-normal density has a mode at 0 which implies that the majority of the producers are clustered near full efficiency level. The assumption may be unnecessarily restrictive particular for industries in which certain degree of inefficiency is expected for the producers. The assumption is relaxed by having $\mu \neq 0$ to allow the mode to depart from 0. Since limited theory is available in guiding the choice of u_i's distribution, various distribution assumptions are explored in the literature for their flexibility in shaping the distribution (e.g., the Gamma distribution of Greene 1980) and/or for checking the robustness of estimation results.

It is often of great empirical interest to estimate the degree of inefficiency (u_i) for each producer (observation). The observation-level estimates are obtained using the estimator $\mathrm{E}(u_i|v_i - u_i)$ proposed by Jondrow et al. (1982). The value of $100 \times \mathrm{E}(u_i|v_i - u_i)$ gives the percentage by which output is increased if production is fully efficient. Similarly, an efficiency index is estimated using $\mathrm{E}(\exp(-u_i)|v_i - u_i)$ (Battese and Coelli 1988). The estimated value gives the actual output as a share of potential output, and the value is bounded between 0 and 1. A likelihood ratio test of the null hypothesis that u_i equals 0 can be performed to test for the presence of inefficiency. It amounts to testing the model against its OLS counterpart (the model without u_i). The distribution of the test statistic, however, is non-standard, because the value of 0 is on the boundary of u_i's support. Alternatively, given that an obvious difference between v_i and $v_i - u_i$ is the skewness of the latter, Schmidt and Lin (1984) suggest a simple test based on the sample skewness of the OLS residuals. If $v_i - u_i$ is the correct specification, the residuals would skew to the left and the null hypothesis of a normal error would be rejected.

If panel data is available, the model may be written as (for the ease of illustration, assume that the deterministic part of the frontier function is linear):

$$\ln y_{it} = \alpha + \boldsymbol{x}_{it}'\boldsymbol{\beta} + v_{it} - u_i, \quad u_i \geq 0, \tag{5}$$

where α is a constant. One may impose distributional assumptions on v_{it} and u_i to derive the likelihood function of the model. Alternatively, a distribution-free approach suggested by Schmidt and Sickles (1984) is available. In this approach, one defines $\alpha_i = \alpha - u_i$ and assumes that the u_i is an individual-specific parameter. With the definition of α_i substituted into (5), the model is estimated by standard fixed-effect panel estimators which yield consistent estimates of α_i for a large T. Since $\alpha_i = \alpha - u_i$ and $u_i \geq 0$, one then recovers the estimated values of α and u_i using the normalization equations of $\hat{\alpha} = \max\{\hat{\alpha}_i\}$ and $\hat{u}_i = \hat{\alpha} - \hat{\alpha}_i$. This normalization procedure amounts to counting the most efficient firm in the sample as 100% efficient.

By duality, technical inefficiency in the production also leads to a higher cost of production. Estimating the cost associated with technical inefficiency is often of important policy values, and it can be done using a stochastic cost frontier model in a cost minimization framework. The model specification is:

$$\ln C_i = \ln C_i^* + u_i', \quad u_i' \geq 0, \tag{6}$$

$$\ln C_i^* = g(\boldsymbol{w}_i, y_i; \boldsymbol{\gamma}) + v_i', \tag{7}$$

where C_i is the observed cost of producer i, C_i^* is the efficient level of cost which is subject to a zero-mean random error v_i', $\boldsymbol{w}_i$ is the vector of input prices, $\boldsymbol{\gamma}$ is the vector of coefficients, and $u_i' \geq 0$ is the effect of inefficiency on the cost of production. Equation (7) defines the stochastic cost frontier, and the observed cost lies above the frontier. The value of $100 * u_i'$ measures the extra cost as a percentage of the minimum cost. Econometric analysis of (6) and (7) is similar to that of the production function model. A notable difference is that the cost model's OLS residuals skew to the right if inefficiency presents in the data.

An advantage of a cost function approach over a production function approach is that the issue of *allocative inefficiency* can be addressed in addition to the technical inefficiency. Allocative inefficiency refers to the use of improper input combinations, i.e. the marginal rate of technical substitution between inputs departs from the input price ratio. The improper input mix increases the cost of production, and the effect is not the same as technical inefficiency. Because the analysis of allocative inefficiency requires information of input prices, it is usually carried out in a cost minimization framework. To jointly estimating both of the technical and allocative inefficiency, Schmidt and Lovell (1979) provide the solution technique for a cost system in which the production technology is Cobb-Douglas. Kumbhakar (1997) presents a theoretical solution for a model with a translog cost function, and the difficulty in the empirical implementation of this model is discussed and resolved in Kumbhakar and Wang (2005b) and Kumbhakar and Tsionas (2005).

Although the stochastic frontier model is most often applied to the estimation of production and cost functions, an increasing body of research has adopted the methodology to other fields of study. Hofler and Murphy (1992) apply this estimation approach to labor market search models. Due to the costs of search, observed wages tend to fall below the maximum offers that are available in the market; this shortfall is analogous to a technical inefficiency. Another application is found in the study of financing constraints on investment, where Wang (2003) models the frictionless level of investment as the frontier, and actual investment falls below the frontier because of financing constraints. This approach allows Wang to quantify the effect of financing constraints on investment (represent by u_i), which is infeasible with the conventional linear-regression approach. In an application to economic growth, Kumbhakar and Wang (2005a) employ the stochastic frontier approach and model growth convergence as countries' movements toward the world production frontier. A country may fall short of producing the maximum possible output because of technical inefficiency, and the phenomenon of technological catch-up is observed if the country moves toward the world production frontier over time. By making u_i a function of time and other macro variables, Kumbhakar and Wang test and confirm the convergence hypothesis.

The stochastic frontier model also finds applications in finance. For example, a long-standing issue in the finance literature is whether the IPO underpricing —an phenomenon that the initial offer price of an IPO is below the closing day's bid price— is deliberate on the firm's part or not. Hunt-McCool et al. (1996) adopt the stochastic frontier model to investigate the issue, in which u_i measures the difference between the maximum predicted offer price and the actual offer price. The advantage of the stochastic frontier model in this application is that it can be used to measure the level of deliberate underpricing in the premarket without using aftermarket information.

Kumbhakar and Lovell (2000) offer an excellent review of the existing models in the stochastic frontier literature. The more recent developments in the literature aim at making the model more flexible. For instance, correlations between v_i and u_i are made possible through copula functions. If time series or panel data are available, then it is possible to make u_t or u_{it} serially correlated. Semi-parametric and non-parametric estimation methods are also adopted to estimate the frontier of the production function (e.g., $f(\boldsymbol{x}_i; \boldsymbol{\beta})$) and the frontier of the cost function (e.g., $g(\boldsymbol{w}_i, y_i; \boldsymbol{\gamma})$) so that they are not restricted to specific functional forms.

References

[1] Aigner, D., Lovell, C.A.K., and Schmidt, P. (1977). "Formulation and Estimation of Stochastic Frontier Production Function Models," *Journal of Econometrics* **6**, pp. 21-37.

[2] Battese, G.E., and Coelli, T.J. (1988). "Prediction of Firm-Level Technical Efficiencies with a Generalized Frontier Production Function and Panel Data," *Journal of Econometrics* **38**, pp. 387-99.

[3] Greene, W.H. (1980). "On the Estimation of a Flexible Frontier Production Model," *Journal of Econometrics* **13**, pp. 101-15.

[4] Hofler, R.A., and Murphy, K.J. (1992). "Underpaid and Overworked: Measuring the Effect of Imperfect Information on Wages," *Economic Inquiry* **30**, pp. 511-29.

[5] Hunt-McCool, J., Koh, S.C., and Francis, B.B. (1996). "Testing for Deliberate Underpricing in the IPO Premarket: A Stochastic Frontier Approach," *Review of Financial Studies* **9**, pp. 1251-69.

[6] Jondrow, J., Lovell, C.A.K., Materov, I.S., and Schmidt, P. (1982). "On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model," *Journal of Econometrics* **19**, pp. 233-38.

[7] Kumbhakar, S.C. (1997). "Modeling Allocative Inefficiency in a Translog Cost Function and Cost Share Equations: An Exact Relationship," *Journal of Econometrics* **76**, pp. 351-56.

[8] Kumbhakar, S.C., and Lovell, C.A.K. (2000). Stochastic Frontier Analysis. Cambridge [England] ; New York: Cambridge University Press.

[9] Kumbhakar, S.C., and Tsionas, E.G. (2005) "The Joint Measurement of Technical and Allocative Inefficiencies: An Application of Bayesian Inference in Nonlinear Random-Effects Models," *Journal of American Statistical Association* **100**, pp. 736-47.

[10] Kumbhakar, S.C., and Wang, H.-J. (2005a). "Estimation of Growth Convergence Using a Stochastic Production Frontier Approach," *Economics Letters*, **3**, pp. 300-305.

[11] Kumbhakar, S.C., and Wang, H.-J. (2005b). "Estimation of Technical and Allocative Inefficiency: A Primal System Approach," *Journal of Econometrics*, *forthcoming.*

[12] Meeusen, W., and van den Broeck, J. (1977). "Technical Efficiency and Dimension of the Firm: Some Results on the Use of Frontier Production Functions," *Empirical Economics* **2**, pp. 109-22.

[13] Schmidt, P., and Lin, T.F. (1984). "Simple Tests of Alternative Specifications in Stochastic Frontier Models," *Journal of Econometrics* **24**, pp. 349-61.

[14] Schmidt, P., and Lovell, C.A.K. (1979). "Estimating Technical and Allocative Inefficiency Relative to Stochastic Production and Cost Frontiers," *Journal of Econometrics* **9**, pp. 343-366.

[15] Schmidt, P., and Sickles, R.C. (1984). "Production Frontiers and Panel Data," *Journal of Business and Economic Statistics* **2**, pp. 367-74.

[16] Wang, H.-J. (2003) "A Stochastic Frontier Analysis of Financing Constraints on Investment: The Case of Financial Liberalization in Taiwan," *Journal of Business and Economic Statistics* **3**, pp. 406-19.