GOODNESS OF FIT TESTS

IN STOCHASTIC FRONTIER MODELS

Wei Siang Wang
Nanyang Technological University

Christine Amsler
Michigan State University

Peter Schmidt
Michigan State University
Yonsei University

In this paper we discuss goodness of fit tests for the distribution of technical inefficiency in stochastic frontier models. If we maintain the hypothesis that the assumed normal distribution for statistical noise is correct, the assumed distribution for technical inefficiency is testable. We show that a goodness of fit test can be based on the distribution of estimated technical efficiency, or equivalently on the distribution of the composed error term. We consider both the Pearson chi-squared test and the Kolmogorov-Smirnov test. We provide simulation results to show the extent to which the tests are reliable in finite samples.

May 2, 2010

1. INTRODUCTION

In this paper we consider the stochastic frontier model introduced by Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977). We write the model as

$$(1) \qquad y_i = X_i\beta + \varepsilon_i \quad , \quad \varepsilon_i = v_i - u_i \quad , \quad u_i \geq 0 \quad , \; \mathrm{i} = 1,\ldots n.$$

Here typically y_i is log output, X_i is a vector of input measures (e.g., log inputs in the Cobb-Douglas case), v_i is a normal error with mean zero and variance σ_v^2, and $u_i \geq 0$ represents technical inefficiency. Technical efficiency is defined as $TE_i = \exp(-u_i)$, and the point of the model is to estimate u_i or TE_i.

A specific distributional assumption on u_i is required. The papers cited above considered the case that u_i is half-normal (that is, it is the absolute value of a normal with mean zero and variance σ_u^2) and also the case that it is exponential. Other distributions proposed in the literature include general truncated normal (Stevenson (1980)) and gamma (Greene (1980a, 1980b, 1990) and Stevenson (1980)). Our exposition is for the cross-sectional case, but we could also consider panel data as in Pitt and Lee (1981).

Our interest is in testing the distributional assumption on u_i. We will do this while maintaining the other assumptions that underlie the model, such as the functional form of the regression, the exogeneity of the X_i, and the normality of v_i. This viewpoint is motivated by the fact that in this literature the specification of the distribution of u_i is often regarded as being subject to the most doubt.

The problem then arises that u_i is not observable, and in fact cannot be consistently estimated. To be more precise, define $\hat{\beta}$ to be the MLE of β and $\hat{\varepsilon}_i = y_i - X_i\hat{\beta}$. Then the usual estimate of u_i, suggested by Jondrow et al. (1982), is $\hat{u}_i = E(u_i | \varepsilon_i)$, evaluated at $\varepsilon_i = \hat{\varepsilon}_i$. The distribution of $\hat{u}_i$ has been derived by Wang and Schmidt (2009). It is not the same as the distribution of u_i, even for large *n*. Therefore it is not legitimate to test goodness of fit by comparing the observed distribution of $\hat{u}$ to the assumed distribution of *u*. It is legitimate to test goodness of fit by comparing the observed distribution of $\hat{u}$ to the distribution derived by Wang and Schmidt. However, it is easier to base the tests instead on the distribution of ε_i that is implied by normality of v_i and the assumed distribution of u_i. This is reasonable because, given that normality of v_i is maintained, a rejection of the implied distribution of ε_i is a rejection of the assumed distribution of u_i.

We consider the usual χ^2 goodness of fit test based on expected and actual numbers of observations in cells, and also the Kolmogorov-Smirnov test based on the maximal difference between the empirical and theoretical cdf. For these tests the only technical problem of note is how to handle the issue of parameter estimation. This is relevant because both the "observations" $\hat{\varepsilon}_i = y_i - X_i\hat{\beta}$ and the expected numbers of observations in various cells (for the χ^2 test) or the theoretical cdf (for the Kolmogorov-Smirnov test) depend on estimated parameters. For the chi-squared test, the relevant asymptotic theory was developed by Heckman (1984), Tauchen (1985) and Newey (1985), and we explain how this theory allows asymptotically valid tests in the stochastic frontier setting. For the Kolmogorov-Smirnov test, the comparable asymptotic theory is given by Bai (2003). The bootstrap can also be used to construct asymptotically valid tests (either for the chi-squared test or for the Kolmogorov-Smirnov test).

The plan of the paper is as follows. In Section 2 we give a brief survey of the literature on specification testing in the stochastic frontier model and we discuss further the basics of goodness of fit testing in this context. Sections 3 and 4 contain a general exposition of goodness of fit tests for simple and composite hypotheses, respectively. Section 5 gives a brief discussion of a prototypical problem, testing for normality, and presents some simulations. In Section 6 we discuss the problem of main interest, testing the error distribution in the stochastic frontier model, and we present detailed simulation evidence on the accuracy (size) and the power of various tests. Section 7 contains two empirical examples. Finally, Section 8 gives our concluding remarks.

2. TESTS BASED ON THE DISTRIBUTION OF ε

There are surprisingly few papers that explicitly address specification testing in stochastic frontier models. Schmidt and Lin (1984) and Coelli (1995) provide tests of the null hypothesis that the composed error is symmetric, which is really the hypothesis that the stochastic frontier model does not apply. Lee (1983) tests the null hypothesis that the distribution of u_i is half-normal (or that it is general truncated normal) against the alternative that its distribution is in a four-parameter Pearson family. He uses the OPG form of the LM test. This is reasonable but it assumes a particular, though flexible, alternative. Kopp and Mullahy (1990) use GMM methods to construct a general specification test. They define a vector of moment conditions based on products of powers of the regressors and powers of the centered (demeaned) residuals. They specify enough moment conditions so that the parameters are overidentified. The correctness of the specification can then be tested using the general test of overidentifying restrictions of Hansen (1982), or alternatively using conditional moment tests of the type suggested by Newey (1985) and Tauchen (1985). (These tests will be discussed in more detail in Section 4 below.) Chen and Wang (2009) is similar in spirit to Kopp and Mullahy, but they make different assumptions about the noise component of the error (v) and they suggest different moment conditions. Bera and Mallick (2002) suggest the information matrix test of White (1982), which can be interpreted as a conditional moment test where the moments that are tested are derived from the information matrix equality.

In this paper we take a possibly more direct and intuitive route and consider goodness of fit tests. For goodness of fit tests, the question is whether some observable quantity does or does not have the distribution that it should have if the model is correct, and then the question arises of what observable quantity to focus on. As noted above, the usual estimate of u_i is $\hat{u}_i = E(u_i|\varepsilon_i)$. (This is evaluated at $\varepsilon_i = \hat{\varepsilon}_i$, a point that we ignore in the rest of this section but address subsequently, when we discuss the relevance of allowing for the effects of parameter estimation.) The distribution of $\hat{u}_i$ is given by Wang and Schmidt (2009). It depends on the assumed distributions for both v_i and u_i, and it is not the same as the distribution of u_i. Therefore it is not legitimate to test goodness of fit by comparing the observed distribution of $\hat{u}$ to the assumed distribution of u. So, for example, if u is assumed to be half-normal, this does not imply that $\hat{u}$ should be half-normal, and it is not correct to test the half-normal assumption by seeing whether the distribution of $\hat{u}$ appears to be half-normal. However, this does not mean that the observed distribution of $\hat{u}$ is uninformative. It is perfectly legitimate to test goodness of fit by comparing the observed distribution of $\hat{u}$ to the distribution that it should have under the distributional assumptions being made, as derived by Wang and Schmidt. Because this distribution depends on the distribution of both v and u, we have to maintain the correctness of the assumed (normal) distribution of v to test the correctness of the distributional assumption on u. This issue is inevitable in this context.

Such a comparison is complicated because the distribution of $\hat{u}$ is complicated. It is much easier to base a goodness of fit test on the distribution of ε. The distribution of ε also follows from the assumed distributions of v and u, and so if we maintain the correctness of the assumed distribution for v, we can test the correctness of the assumed distribution for u by a goodness of fit test based on the distribution of ε. This is computationally easier than a test based on the distribution of $\hat{u}$. The following simple point is therefore relevant: $\hat{u}$ is a monotonic function of ε. This implies that most goodness of fit tests based on the distribution of $\hat{u}$ will be equivalent to the same goodness of fit tests based on the distribution of ε. For example, the Kolmogorov-Smirmov test will be exactly the same whether it is based on the distribution of $\hat{u}$ or the distribution of ε. For the Pearson χ^2 test based on the observed versus actual numbers of observations in cells, again the test is exactly the same whether it is based on the distribution of $\hat{u}$ or the distribution of ε, provided that the cells are defined conformably.

Therefore, for reasons of computational simplicity, we will consider tests based on the distribution of ε that is implied by the assumed distributions for v and u. We maintain the correctness of the assumed (normal) distribution of v, and therefore interpret the tests as tests of the correctness of the assumed distribution of u.

3. SIMPLE HYPOTHESES

Suppose that we have a random sample $y_1, y_2, ..., y_n$ and we wish to test the hypothesis that the population distribution is characterized by the pdf $f(y, \theta_0)$. The subscript "zero" on θ indicates the true value of the parameter θ, which we assume to be the same as the value specified by the hypothesis being tested. That is, in this section we take θ_0 as given. Thus, for example, we could be testing the simple hypothesis that *y* is distributed as N(0,1), as opposed to the composite hypothesis that *y* is normal with μ and σ^2 unspecified.

To define the Kolmogorov-Smirnov statistic, let $F(y,\theta_0)$ be the cdf corresponding to the pdf $f(y,\theta_0)$. Also let $F_n(y)$ be the empirical cdf of the sample: $F_n(y) =$ (number of $y_i \le y$)/*n*. Then the Kolmogorov-Smirnov statistic is

(2) $$\text{KS} = \sup_y \left| F(y,\theta_0) - F_n(y) \right|$$

The asymptotic distribution of KS is known and widely tabulated. It does not depend on the form of the distribution (*f* or *F*).

Now consider the Pearson χ^2 statistic. Let the possible range of *y* be split into *k* "cells" (intervals) $A_1,\ldots,A_k$, such that any value of *y* is in one and only one cell. Let $1(y \in A_j)$ be the "indicator function" that equals one if y is in cell A_j, and equals zero otherwise. Let $p_j = p_j(\theta_0) = P(y \in A_j) = E[1(y \in A_j)]$. With *n* observations as above, we define the observed (*O*) and expected (*E*) numbers of observations in each cell:

(3) $$O_j = \sum_{i=1}^{n} 1(y_i \in A_j) \quad , \quad E_j = np_j \quad , \quad j = 1,\ldots,k .$$

Then the Pearson χ^2 statistic is given by:

(4) $$\chi^2 = \sum_{j=1}^{k} (O_j - E_j)^2 / E_j$$

Asymptotically (as $n \to \infty$) its distribution is chi-squared with (*k*-1) degrees of freedom.

It is interesting (and later it is useful) to put these results into a generalized method of moments (GMM) framework. We begin with the set of moment conditions

(5) $$E[g(y,\theta_0)] = 0$$

where $g(y, \theta)$ is a vector of dimension (k-1) whose j^{th} element equals $[1(y \in A_j) - p_j(\theta)]$. The subscript "zero" on θ reinforces the point that the expectation in (5) equals zero only at θ_0, the true value of θ. Also, note that we have omitted one cell so as to avoid a subsequent singularity. We have omitted cell A_k but the choice of which cell to omit does not matter. Now define

(6) $$\bar{g}(\theta) = \frac{1}{n}\sum_{i=1}^{n} g(y_i, \theta)$$

and note that the j^{th} element of $\bar{g}(\theta)$ is equal to $\frac{1}{n}(O_j - E_j(\theta))$. We also need to define the variance matrix of the moment conditions $g(y, \theta_0)$. This variance matrix is the matrix $V(\theta_0)$, of dimension (k-1) by (k-1), whose j^{th} diagonal element equals ($p_j - p_j^2$), and whose i, j^{th} off diagonal element ($i \neq j$) equals ($-p_i p_j$), with all probabilities evaluated at θ_0.

A central limit theorem implies that the asymptotic distribution of $\sqrt{n}\bar{g}(\theta_0)$ is N($0, V(\theta_0)$). From this fact it follows that

(7) $$n\bar{g}(\theta_0)'V(\theta_0)^{-1}\bar{g}(\theta_0) \rightarrow_d \chi^2_{k-1} \ .$$

To link this to the distributional result given above for the test of the simple hypothesis that y has density $f(y, \theta_0)$, we simply observe that

(8) $$n\bar{g}(\theta_0)'V(\theta_0)^{-1}\bar{g}(\theta_0) = \sum_{j=1}^{k}(O_j - E_j)^2 / E_j$$

which is the Pearson χ^2 statistic. The equality in (8) is proved in Appendix A. So this establishes the distributional result given in the sentence following equation (4).

There seems to be a general consensus in the literature that the Kolmogorov-Smirnov test should be more powerful than the Pearson χ^2 test. The Pearson test is based on groupings (cells) that are inevitably somewhat arbitrary (except perhaps where the variable of interest is discrete).

More formally, the Kolmogorov-Smirnov test is consistent against essentially all alternatives, whereas the Pearson test is inconsistent against alternatives that imply the same cell probabilities as under the null. Whether such arguments are relevant in finite samples is a question that we address in our simulations.

4. COMPOSITE HYPOTHESES

Now suppose that we wish to test the composite hypothesis that the population distribution is characterized by the pdf $f(y,\theta)$ for some (unspecified and unknown) value of θ. This is the empirically relevant case.

For the Kolmogorov-Smirnov test, we can estimate θ by MLE. Let this estimate be $\hat{\theta}$. Now we can use $\hat{\theta}$ in place of θ_0 in equation (2) to calculate the statistic. The problem is that the distribution of the statistic is changed, even asymptotically, and this change depends in general on both the distribution of the data and the value of θ_0. Bai (2003) shows how to modify the statistic so that the modified statistic has an asymptotic distribution that does not depend on the distribution of the data or the value of θ_0. Therefore critical values can be tabulated. He uses results from Khmalzade (1981, 1988, 1993) on transformation of the empirical process into a martingale. These are conceptually difficult papers, especially in the case that θ_0 is multi-dimensional (which it is in the stochastic frontier model), though the required computations are not difficult.

An asymptotically valid Kolmogorov-Smirnov test for a composite hypothesis can also be constructed using bootstrapping. Let $f(y,\hat{\theta})$ be the hypothesized density, evaluated at the estimate $\hat{\theta}$. Now we use a "parametric bootstrap": for $b = 1, 2, \ldots, B$, where B (the number of bootstrap draws) is large, draw $y_1^{(b)}, y_2^{(b)}, \ldots, y_n^{(b)}$ from $f(y,\hat{\theta})$. Based on this data, calculate the estimate $\hat{\theta}^{(b)}$ and the KS statistic in (2) based on $\hat{\theta}^{(b)}$. Then use the critical values derived from the appropriate quantiles of the empirical distribution of these *B* values of the statistic. The asymptotic validity of this procedure has been established by Giné and Zinn (1990) and Stute, Gonzáles and Presedo (1993).

For the stochastic frontier model given in equation (1) above, the discussion of the previous paragraph does not precisely apply, because we are interested in testing a hypothesis about the distribution of an unobservable (ε_i). Nevertheless, ε_i is a function of data and parameters, and the bootstrap, if properly applied, should properly account for the variability due to the fact that it must be evaluated using estimated parameters. To be explicit about the bootstrap procedure that is used in this case, let $\theta = (\beta', \sigma_u^2, \sigma_v^2)'$ and let $\hat{\theta}$ be the MLE from the original data. This leads to residuals $\hat{\varepsilon}_i$ and a *KS* statistic that compares the empirical cdf of the $\hat{\varepsilon}_i$ to the theoretical cdf implied by $\hat{\theta}$. Now, for bootstrap draw *b*, we draw errors $\varepsilon_i^{(b)}$ ($i = 1,\ldots,n$) from the composed error distribution implied by $\hat{\theta}$, and we generate the bootstrap data $y_i^{(b)} = X_i\hat{\beta} + \varepsilon_i^{(b)}$. From the bootstrap data we obtain new estimates $\hat{\theta}^{(b)}$ and residuals $\hat{\varepsilon}_i^{(b)}$, and a Kolmogorov-Smirnov statistic $KS^{(b)}$ that compares the empirical cdf of the $\hat{\varepsilon}_i^{(b)}$ to the theoretical cdf implied by $\hat{\theta}^{(b)}$. Then the critical values for the (original) *KS* statistic are the appropriate quantiles of the empirical distribution of the *B* values of $KS^{(b)}$. So although the *KS* statistic based on the original data is based on $\hat{\varepsilon}_i$ rather than ε_i, the bootstrap procedure copies this distinction exactly in the bootstrap samples, since $KS^{(b)}$ is based on the $\hat{\varepsilon}_i^{(b)}$, not the $\varepsilon_i^{(b)}$.

Next we will consider the Pearson χ^2 test. As for the Kolmogorov-Smirnov test, it is not legitimate to ignore parameter estimation. Also as for the Kolmogorov-Smirnov test, an asymptotically valid test can be obtained using critical values from the parametric bootstrap. However, for the χ^2 test the necessary asymptotic theory to correct for parameter estimation is relatively straightforward, and tests based on this theory are a computationally simple alternative to tests using the bootstrap.

To discuss this asymptotic theory, recall that the number of cells was *k*, and let the dimension of θ be *m*, with $m \leq k-1$. We have $p_j = p_j(\theta)$ and $E_j(\theta) = np_j(\theta)$; that is, the expected numbers of observations in the cells depend on θ. Thus the value of the statistic in (4) or (8), say $\chi^2(\theta)$, depends on θ. As above, let $\hat{\theta}$ be the MLE of θ, and let $\tilde{\theta}$ be the value of θ that minimizes $\chi^2(\theta)$. A famous result is that $\chi^2(\tilde{\theta})$ is asymptotically distributed as chi-squared with (*k*-1-*m*) degrees of freedom. That is, we still have a chi-squared distribution but the number of degrees of freedom is reduced by one for every estimated parameter. This is a nice result but it is not altogether satisfying, since we have reduced the number of degrees of freedom, and because $\tilde{\theta}$ is in general an inefficient estimator. It is much more natural to consider the statistic $\chi^2(\hat{\theta})$ which uses the MLE. However, this is not asymptotically distributed as chi-squared, and using the chi-squared distribution with (*k*-1-*m*) degrees of freedom results in a test which is conservative, and therefore presumably less powerful than is possible. (For this result, and the result referred to above as famous, see, e.g., Tallis (1983), p. 457.)

To understand these results, and the way in which parameter estimation by MLE is successfully accommodated, we return to the GMM interpretation of the χ^2 test given at the end of Section 3. The value of θ is unknown, but we can estimate θ by GMM based on the moment conditions (5). In this case we will minimize the GMM criterion function

(9) $$n\bar{g}(\theta)'\hat{V}^{-1}\bar{g}(\theta)$$

where $\hat{V}$ is either $V(\theta)$, in the case of the "continuous updating" GMM estimator, or is any consistent estimate of $V(\theta_0)$, in the case of the "two step" GMM estimator. The first possibility corresponds to the minimization of $\chi^2(\theta)$ with respect to θ, and yields the estimator $\tilde{\theta}$ discussed in the previous paragraph. In either the continuous updating case or the two step case, standard GMM results indicate that the minimized value of the criterion function (9) is asymptotically distributed as chi-squared with degrees of freedom equal to the number of moment conditions minus the number of parameters estimated, that is, (k-1-m) degrees of freedom. (This is generally referred to in the GMM literature as the "test of overidentifying restrictions.") This argument establishes the "famous result" referred to above.

The estimator $\tilde{\theta}$ is not generally efficient, and so we ought to be able to do better than this. As above, let $\hat{\theta}$ be the MLE, which is (asymptotically) efficient. Unfortunately $\chi^2(\hat{\theta})$ does not generally have a chi-squared distribution. This raises the question of whether we can construct a goodness of fit statistic based on $\hat{\theta}$ that does have a chi-squared distribution. The answer is yes, as was shown by Heckman (1984), Tauchen (1985) and Newey (1985). Our discussion will follow Tauchen. We wish to test the composite hypothesis that the density of y is $f(y,\theta)$. Define the "score function"

(10) $$s(y,\theta) = \frac{\partial \ln f(y,\theta)}{\partial \theta} .$$

The MLE satisfies the first order condition $\sum_{i=1}^{n} s(y_i,\hat{\theta}) = 0$ and is the GMM estimator based on the (exactly identified) set of moment conditions: $Es(y,\theta_0) = 0$. Now the technical trick that leads to the test is to combine these moment conditions based on the score function with the moment conditions that we want to test, based on numbers of observations falling into various cells. Formally we write the full set of moment conditions as $Eh(y,\theta_0)=0$, where

(11) $$h(y,\theta)=\begin{bmatrix} h_1(y,\theta) \\ h_2(y,\theta) \end{bmatrix}=\begin{bmatrix} s(y,\theta) \\ g(y,\theta) \end{bmatrix}.$$

Here $h_1=s$ is the score function and $h_2=g$ is the vector of (k-1) functions given in equation (5) above. We wish to maintain the correctness of h_1 (to obtain $\hat{\theta}$) and test the correctness of h_2.

The test statistic is of the form

(12) $$\text{CMT} = n\bar{g}(\hat{\theta})'\hat{C}^{22}\bar{g}(\hat{\theta})$$

where $\hat{C}^{22}$ will be defined in the next paragraph. The relevant distributional result is that CMT is asymptotically distributed as chi-squared with (k-1) degrees of freedom. That is, we do obtain a chi-squared limiting distribution *and* there is no loss in degrees of freedom due to estimation of θ.

The difference between this statistic and $\chi^2(\hat{\theta})$ is that the conditional moment test (CMT) uses $\hat{C}^{22}$ where $\chi^2(\hat{\theta})$ uses $V(\hat{\theta})^{-1}$, where $V(\theta)$ is the variance matrix of $g(y,\theta)$ The matrix $\hat{C}^{22}$ is defined as follows. Let C be the variance matrix of the vector $h(y,\theta)$. Its dimension is (m+k-1) by (m+k-1). Let C^{-1} be its inverse. We partition C and C^{-1} correspondingly to the partitioning of h into h_1 and h_2:

(13) $$C=\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}, \quad C^{-1}=\begin{bmatrix} C^{11} & C^{12} \\ C^{21} & C^{22} \end{bmatrix}.$$

So C^{22} is the lower right submatrix, of dimension (k-1) by (k-1), of C^{-1}. Then $\hat{C}^{22}$ is any consistent estimate of C^{22}. (A specific estimate will be discussed below.) We can note that $C_{22}=V(\theta)$ is the variance matrix of $g(y,\theta)$ and so basically $\chi^2(\hat{\theta})$ uses (an estimate of) C_{22}^{-1} whereas CMT uses (an estimate of) C^{22}. A standard matrix equality says that

(14) $$C^{22} = (C_{22} - C_{21}C_{11}^{-1}C_{12})^{-1}$$

which is bigger than C_{22}^{-1}. That is the sense in which the CMT adjusts for the fact that $\chi^2(\hat{\theta})$ is too conservative.

Equation (14) shows that an estimate of C^{22} requires an estimate of all of C. The most commonly used estimate is the "OPG" (for "outer product of the gradient") estimate:

(15) $$\hat{C} = \hat{C}(\hat{\theta}) = \frac{1}{n}\sum_{i=1}^{n} h(y_i, \hat{\theta})h(y_i, \hat{\theta})'.$$

The CMT using this estimate can be calculated as the uncentered R^2 in a regression of a constant (one) on $h(y_i, \hat{\theta})$. See Newey (1985), p. 1052, for this expression.

Alternatively, for some problems we may be able to obtain an analytical expression for *C,* and then evaluate it at $\hat{\theta}$. The submatrix C_{11} is the information matrix for the estimation of θ by MLE, and may be evaluated using expectations of second derivatives or cross products of first derivatives. The submatrix C_{22} can be evaluated as $V(\hat{\theta})$ where the matrix *V* is defined in the discussion following equation (6). This leaves $C_{12} = Eh_1h_2'$ for which there may be an analytical expression in certain simple cases (e.g. testing for normality), but not in general.

An alternative to the CMT test is to estimate θ by GMM using the full set of moment conditions *h*, and then perform the usual GMM overidentification test. Let $\breve{\theta}$ be this estimate. It is different from the MLE $\hat{\theta}$, but it has the same asymptotic distribution since the second set of moment conditions (*g*) is statistically redundant for estimation given the score (*s*). The GMM overidentification test statistic is $n\bar{h}(\breve{\theta})'\hat{C}^{-1}\bar{h}(\breve{\theta})$. This can be compared to the CMT statistic $n\bar{g}(\hat{\theta})'\hat{C}^{22}\bar{g}(\hat{\theta}) = n\bar{h}(\hat{\theta})'\hat{C}^{-1}\bar{h}(\hat{\theta})$, where the last equality follows from the fact that $\bar{s}(\hat{\theta}) = 0$. So the difference between the CMT and the overidentification test is just due to the difference between $\hat{\theta}$ and $\breve{\theta}$, which is asymptotically negligible. The reason the CMT is preferable is a matter of simplicity – if one has estimated the model by MLE, which would typically be the case, then there is no need to reestimate the model.

There is a further theoretical point worth making. The discussion above takes the cells (A_j) as given, so that the probabilities of observations falling into the cells depend on θ. This fits well into the GMM setting because the usual asymptotic distribution theory for GMM depends on the moment conditions being differentiable with respect to θ. In practice, however, the cell definitions will naturally depend on θ. For example, if we test normality based on five equi-probable cells, the first cell will be $(-\infty, \mu - 0.84\sigma]$. So then the probabilities of being in the various cells are given, but the observed numbers in the cells depend on μ and σ. Now the moment conditions depend on the indicators $1(y \in A_j(\theta))$ which are not differentiable with respect to θ. However, Tauchen shows that the distributional theory for the CMT still holds in this case.

An interesting theoretical detail that we have not found in the literature is the following. Instead of comparing observed and expected numbers of observations in cells, we could compare the sample and population quantiles. For example, in the normal example mentioned above, in the first case we fix the cell boundary at $(\hat{\mu} - 0.84\hat{\sigma})$ and see how close the number of observations less than this is to 0.2. Alternatively, we could calculate the 20$^{\text{th}}$ percentile in the sample and see how close it is to $(\hat{\mu} - 0.84\hat{\sigma})$. Intuition would suggest that the difference between these two tests ought to be asymptotically negligible. The sense in which that is true is discussed in Appendix B.

A final point is that it is also possible to apply the bootstrap to the asymptotically valid (Tauchen) form of the χ^2 statistic. It is well known that the bootstrap can provide higher-order refinements to the asymptotically valid distributions of estimators or test statistics. That is, using asymptotic theory plus the bootstrap may give better (more accurate in finite samples) critical values than using asymptotic theory alone or the bootstrap alone. We will investigate that issue in our simulations.

5. AN INTRODUCTORY EXAMPLE: TESTING NORMALITY

Our interest in this paper is testing distributional assumptions in stochastic frontier models. However, first we will present a few simulations for a simpler problem, testing normality. The point is to see how the tests work in a very simple setting, and in particular in one where we can do some things analytically that we cannot do in the more complicated model.

All of our tests will be based on standard normal data. The number of replications in the simulations is 10,000, except that for the bootstrap tests we use 1000 replications and 999 bootstrap samples per replication. We will consider sample sizes n = 50, 100, 250 and 500. We use tests with nominal size of 0.05. (We also calculated results for sizes of 0.01 and 0.10 but they would not lead to different conclusions than the results for size of 0.05.) For the χ^2 tests we will consider three different numbers of cells: k = 3, 5 and 10. In all cases we use equiprobable cells.

We first consider the test of the simple hypothesis that the data are N(0,1). That is, the values $\mu_0 = 0$ and $\sigma_0 = 1$ are specified by the null hypothesis (i.e. assumed known). The χ^2 statistic is $n\bar{g}(\theta_0)'V(\theta_0)^{-1}\bar{g}(\theta_0)$ as given in equation (8) above, while the KS statistic is as given in equation (2). The results for these tests are given in Table 1A, using the asymptotic critical values and the bootstrapped critical values. They are very easy to summarize. All of the tests are quite accurate, in the sense that actual size is quite close to nominal size. The KS test is very slightly undersized for the smaller sample sizes (n = 50 and 100) but this is not a serious discrepancy. Bootstrapping fixes this problem.

Next we consider the test of the composite hypothesis that the data are N(μ,σ^2) for unknown μ and σ^2. In Table 1B we give the empirical size of four of the asymptotically valid tests discussed above. The first (under the heading "Pearson (Tauchen)") is the Tauchen version of the χ^2 statistic, equal to $n\overline{g}(\hat{\theta})'\hat{C}^{22}\overline{g}(\hat{\theta})$, as given in equation (12). The MLE of θ is $\hat{\theta}=(\hat{\mu},\hat{\sigma}^2)$ where $\hat{\mu}=\overline{y}$ and $\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(y_i-\overline{y})^2$. Then $\hat{C}^{22}$ is obtained from the inverse of $\hat{C}$, the OPG estimate as in equation (15), with $h(y,\theta)$ evaluated at the MLE $\hat{\theta}$. This expression requires the score function for the normal density, which is given in Appendix C. The second test ("Bootstrap Pearson (Tauchen)") uses the same test statistic but uses bootstrapped critical values. The third test ("Bootstrap Pearson") is the usual Pearson χ^2 test given in (8), using $\hat{\theta}$ in place of θ_0, with bootstrapped critical values. The fourth test ("Bootstrap KS") is the Kolmogorov-Smirnov test given in (2), but using $\hat{\theta}$ in place of θ_0, with bootstrapped critical values.

The results for these four tests in Table 1B are easy to summarize. All of the tests that use critical values from the bootstrap are quite accurate, in the sense that empirical size is close to nominal size. The results for the Pearson (Tauchen) test, which relies on asymptotic theory instead of the bootstrap, are less favorable. For this test there are moderate to large size distortions in small samples, especially when a large number of cells is used. For example, with k = 3 the actual size of the nominal 0.05 level test is 0.067 for n = 50 and 0.055 for n = 100, which is not too bad. For k = 5, we have actual sizes of 0.086, 0.070 and 0.057 for n = 50, 100 and 250, respectively, so the size distortions disappear more slowly. For k = 10, we have actual sizes of 0.185, 0.117, 0.079 and 0.060 for n = 50, 100, 250 and 500, respectively, so that a rather large sample size is required to have an accurate test. An obvious implication is not to use too many cells unless the sample size is large (or, to use the bootstrap).

It is interesting to ask why it is that this test is less accurate than the Pearson test for the simple hypothesis (Table 1A). The current test differs from the Pearson test of the simple hypothesis in a number of ways. First, it uses cells defined on the basis of $\hat{\theta}$ rather than θ_0. Second, it obtains the weighting matrix as a submatrix of the variance matrix of the moment conditions after they have been augmented with the score. Third, it evaluates the variance matrix of the (augmented) moment conditions using the OPG estimate, as opposed to using an analytical expression.

The question is which of these differences is the one that matters. To provide evidence on this question, we consider two other tests. One ("Simple Hypothesis OPG" in Table 1B) is based on the test statistic for the simple hypothesis (equation (8) above) except that we replace the variance matrix $V(\theta_0)$ with the OPG estimate $\hat{C}_{22}(\theta_0) = \frac{1}{n}\sum_{i=1}^{n} g(y_i, \theta_0) g(y_i, \theta_0)'$. So we are using the OPG estimate unnecessarily, but we do not have estimation error in θ. The second test ("Composite Hypothesis *C* Known" in Table 1B) is based on the test stastistic $n\bar{g}(\hat{\theta})' C^{22}(\theta_0) \bar{g}(\hat{\theta})$, where C^{22} is the relevant submatrix of the true (not estimated) variance matrix *C* of the augmented moment conditions. This is possible because for this problem we can calculate *C* analytically. This calculation is given in Appendix C. This test statistic uses the cell definitions based on $\hat{\theta}$ but uses the analytical expression for *C* and evaluates it at θ_0. This would be infeasible in actual practice but it is feasible in the Monte Carlo setting because we know the value of θ_0.

The results for "Composite Hypothesis *C* Known" are quite good. The results for "Simple Hypothesis OPG" are much less accurate. The size distortions follow the same pattern as for "Pearson (Tauchen)" though they are a little smaller. Thus it appears that the finite sample inaccuracy of the non-bootstrapped Tauchen version of the Pearson test is due primarily to the use of the OPG estimate of *C*. The true value of *C* accounts properly for the effects of parameter estimation but the OPG estimated $\hat{C}$ does not.

6. THE STOCHASTIC FRONTIER MODEL

We now return to the stochastic frontier model. The model is given in equation (1) above. We will consider first, and in most detail, the case in which v is normal and u is half-normal. The parameters of the model are β, σ_v^2 and σ_u^2. We follow the usual notational convention that σ_u^2 is the variance of the normal random variable of which u is the absolute value, so that $\text{var}(u) = \frac{\pi - 2}{\pi}\sigma_u^2$. We also adopt the standard notation $\lambda = \sigma_u / \sigma_v$ and $\sigma^2 = \sigma_u^2 + \sigma_v^2$.

We will refer to $\varepsilon = v - u$ as the "composed error". Its density is known (e.g. Aigner, Lovell and Schmidt (1977), p. 26) but its cdf does not have any known closed-form expression. We need the cdf, or a tabulation of it, to calculate the cell probabilities or the cell boundaries for the χ^2 test, or to calculate the theoretical cdf for the Kolmogorov-Smirnov test. We have therefore tabulated the cdf via a simulation, with 10,000,000 replications for each simulation. In Table 2 we present the quantiles 0.1, 0.2, …, 0.9 for values of λ between zero and 10,000, and for $\sigma^2 = 1$. For a given value of λ and for a different value of σ^2, one just needs to multiply the quantile by σ. (For example, for $\lambda = 1$ and $\sigma^2 = 1$, the 0.40 quantile is -0.754. So for $\lambda = 1$ and $\sigma^2 = 2$, the 0.40 quantile is (-0.754)($\sqrt{2}$) = -1.066.) For values of λ not in the table, interpolation will be needed. A much more detailed set of tables, which gives quantiles from 0.01 to 0.99, is available as an electronic file, on request from the authors.

6.1 Size of the Test

Primarily to check these tabulations, we first briefly consider the case in which the composed error ε is observed (equivalently, β is known or specified by the null hypothesis), and the null hypothesis specifies the value of σ_v^2 and σ_u^2. Thus we are testing a simple null hypothesis under which the distribution of ε is completely specified. We note that, apart from the randomness of the simulation, the results should be exactly the same as the results in Table 1A, where we were testing the null that the data are standard normal. (If the null completely and correctly specifies the distribution of the data, the test is the same as if we compared the percentiles corresponding to the observations to the uniform distribution, and the nature of the parent distribution is irrelevant.) We give results for the composed error with $\sigma_u^2 = \sigma_v^2 = 1$ in Supplemental Table 1, the first table of a supplemental set of tables available from the authors on request. This table corresponds to Table 1A of this paper for the standard normal case. The results are so nearly the same as those in Table 1A that we need not display them here.

Now we turn to the case of main interest, in which we wish to test that the composed error has the normal / half-normal distribution with unspecified (unknown) values of the parameters. In this case our model for the simulations will be:

(16) $$y_i = \alpha + \varepsilon_i \quad , \quad \varepsilon_i = v_i - u_i$$

and the unknown parameters are α, σ_u^2 and σ_v^2. There are no regressors other than intercept, and no slope coefficients, because estimation error in these is not likely to be important. (It has been observed in many empirical studies that different distributional assumptions do not much change the estimated slope coefficients, whereas the estimated intercepts do change.) However, the presence of intercept is important. It is obviously empirically relevant, and it affects the results because, by demeaning the data, it prevents the level of the series from containing information about σ_u^2. In performing our experiments, we parameterize in terms of $\alpha, \lambda = \sigma_u / \sigma_v$ and σ_v^2. The results depend only on λ but in estimation we estimate three parameters. (For a given value of λ, changing α and σ_v^2 just changes location and scale, and results in a linear transformation of the data which does not change any of the test statistics.)

A technical complication worth mentioning is the "wrong skew" problem pointed out by Waldman (1982). The distribution of ε has a negative skew (negative third central moment). However, in the data we may encounter a positive skew of the residuals. In our case, this would correspond to $m_3 > 0$ where $m_3 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^3$, an occurrence of the "wrong skew." When we have the wrong skew, the MLE's are as follows:

$$(17) \qquad \hat{\alpha} = \bar{y}, \hat{\sigma}_u^2 = 0, \hat{\sigma}_v^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2 .$$

This happens with a positive probability that depends on λ and n. For example, when λ is near zero and n is small, the wrong skew problem occurs nearly half of the time. It is widely argued (e.g. Simar and Wilson (2010)) that the wrong skew problem causes considerable difficulties in inference in stochastic frontier models. One of the points of our experiments will be to see whether this is true for goodness of fit testing.

Our results for the cases where the null is true are given in Tables 3A-3E, which correspond to $\lambda = 0.1$, 0.5, 1, 2 and 10, respectively. These tables have essentially the same format as Table 1B (minus its last two columns), except that they also report the frequency of occurrence of the wrong skew problem.

One striking result in these tables is that the frequency of rejection (size of the test) does not depend very strongly on λ. That is, for a given value of n and for a given test, the size of the test is approximately the same in all five tables. In fact, the results in these tables are very similar to the results in Table 1B, which was for the case of testing normality with unknown mean and variance. The parameter estimation problem is much simpler in the normal case than in the normal / half-normal case, so we might expect larger size distortions in Tables 3A-3E than in Table 1B. However, we don't actually find that; any differences are very slight. Correspondingly, the main conclusions are the same as in Section 5. All of the tests that use bootstrapped critical values are quite accurate (size close to nominal size). The Tauchen version of the Pearson test, which relies on asymptotic theory instead of the bootstrap, is less reliable. There are noticeable size distortions unless the sample size is very large or the number of cells used is small. Based on these results we would recommend using critical values from the bootstrap. The choice of which test to use logically would depend on considerations of power, which we will discuss in the next subsection.

The frequencies of occurrence of the wrong skew problem are in line with previous evaluations, such as in Simar and Wilson (2010). The fact that the frequency of occurrence of the wrong skew problem varies strongly with λ, but the size of the test does not, would seem to imply that any size distortions we encounter are not primarily a reflection of this problem. As a matter of curiosity, we also calculated the frequency of rejection for those samples where the wrong skew problem did and did not occur. We did this for the Tauchen version of the Pearson test only, since that was the only test with significant size distortions. These results are given in Supplemental Table 2, in the supplemental set of tables available from the authors on request. The frequencies of rejection are different but not too different for the samples with the wrong skew than they are for the samples with the correct skew. For example, with $n = 50$ and $\lambda = 1$, we have 6399 replications with the correct (negative) skew and 3601 with the wrong (positive) skew. For the Pearson (Tauchen) χ^2 test with $k = 3$, we have rejection frequencies of 0.094 conditional on correct skew and 0.055 conditional on wrong skew; for $k = 5$ we have 0.104 and 0.074. These numbers are clearly different, but it is not the case that the rejections are coming overwhelmingly from one case or the other.

6.2 Power of the Test

Now we turn to the question of the power of the test. This requires specification of the alternative hypothesis. The null is exactly as in the previous section: the model is as given in equation (16), and the null is that the composed error $\varepsilon = v - u$ has the distribution implied by v being normal and u being half-normal. The alternatives that we consider will be based on the same model, except that u will follow some other one-sided distribution. Specifically, we will consider exponential and gamma distributions for u.

For the simulations in this subsection, we still use 10,000 replications for the Tauchen version of the Pearson test, and 1000 replications with 999 bootstrap samples for the tests based on the bootstrap, except that for the bootstrapped KS test, we use 1000 replications with 399 bootstrap samples.

Tables 4A, 4B and 4C give the power of the test when v is N(0,1) and u is exponential with mean equal to θ (and, correspondingly, variance equal to θ^2). We consider $\theta = 0.1$, 0.5, 1, 2, 5 and 10. Varying θ changes the relative importance of noise and one-sided error. Since the results of the tests are invariant to linear transformation of the data, we could equally have kept θ fixed and changed the variance of v. (For example, the results with $\theta = 5$ and $\sigma_v^2 = 1$ are the same as with $\theta = 1$ and $\sigma_v^2 = 1/25$.) Larger values of θ correspond to less noise relative to one-sided error, and presumably should lead to higher power, since it is easier to distinguish half-normal and exponential data if they are contaminated with less noise. As a result, as we move down in each section of these tables, power should increase as θ increases. However, it is not the case that the power goes to one as $\theta \to \infty$. Rather, as $\theta \to \infty$, power should approach the power that we would have if there were no noise and we were testing the null that the data are half-normal against the alternative that they are exponential.

The KS test using bootstrapped critical values is generally the most powerful test. It clearly dominates the other two tests that use bootstrapped critical values. Its comparison to the non-bootstrapped Pearson (Tauchen) test is somewhat ambiguous, because the Pearson (Tauchen) test sometimes appears to be more powerful, but this occurs in cases (small n and/or large k) in which the Pearson (Tauchen) test had non-trivial size distortions. Even in those cases the bootstrapped KS test is more powerful if θ is large enough that power is non-trivial. Basically whenever power is over 0.2, the bootstrapped KS test is best.

Comparing results for the various Pearson tests across the three tables, we see that power is generally higher when less cells are used. That is, power is higher with three cells than with five, and higher with five cells than with ten. (There are a few exceptions for the non-bootstrapped Pearson (Tauchen) test when power is low and n and k are such that size distortions were found under the null.) Since size distortions are smaller and power is higher when a small number of cells is used, it is obvious to recommend using a small number of cells. Precisely how small is a question that could be investigated further.

Unfortunately, we can also see that power is rather low unless the sample size is quite large and/or the variance of u is much larger than the variance of v. For example, when $\theta = 1$, which corresponds to equal variance for v and u, power for the bootstrapped KS test is 0.054 for $n = 50$, 0.089 for $n = 100$, and 0.150 for $n = 250$. When θ is bigger the situation is more favorable. For example, when $\theta = 5$, power is 0.530 for $n = 100$ and 0.930 for $n = 250$. However, $\theta = 5$ corresponds to var(u) = 25var(v), which is not common in empirical applications.

Another way to summarize these results is that we can expect to distinguish exponential data from half-normal, but that this becomes difficult if the data are contaminated by normal noise.

In Tables 5A-5D we consider the case that the one-sided error has a gamma distribution. Now $u = cu^*$ where u^* follows the standard gamma distribution with density

$$f(u^*) = \frac{(u^*)^{m-1} e^{-u^*}}{\Gamma(m)} \quad . \tag{18}$$

The parameter m governs the shape of u^*. When $m = 1$ we have the exponential distribution which we have just considered. Values of m less than one lead to densities with a mode at zero and very steep decline as u^* increases. Values of m larger than one lead to a positive mode, and the distribution approaches normality as $m \to \infty$. We consider $m = 0.1$, 0.5, 2 and 10. The mean and variance of the standard gamma distribution in (18) both equal m, so the mean of $u = cu^*$ equals cm and the variance equals $c^2 m$. Thus for a given value of m, we expect power to increase when c increases.

In Tables 5A-5D, the results for the Pearson tests are for $k = 3$ only.

The general pattern of results is similar to what was found for the exponential case. Power increases as n increases and as c increases. The Kolmogorov-Smirnov test is generally the most powerful. And, again, the power of the tests is low over the part of the parameter space that would seem to be empirically most common.

An interesting feature of these results is that the power is quite low for the values of m greater than one, even for large values of c. This is so despite the fact that the gamma distribution with m greater than one does not at all resemble the half-normal distribution. The reason for this low power is presumably that the gamma distribution with large m resembles the normal distribution, and therefore is mistaken for part of the noise.

6.3 The Exponential Case

In Sections 6.1 and 6.2 the null hypothesis was that the composed error had the normal / half-normal distribution. In this Section we consider the case that the null hypothesis is that the composed error has the normal / exponential distribution. More explicitly, under the null hypothesis $\varepsilon = v - u$ where v is $N(0, \sigma_v^2)$ and u is exponential with mean equal to θ.

In Table 6 we present selected quantiles of the distribution of $\varepsilon = v - u$ when v is normal and u is exponential. It has the same format as Table 2 did for the normal / half-normal case. To display the results we define $\lambda = \theta / \sigma_v$ and $\sigma^2 = \theta^2 + \sigma_v^2$. The tabulations are for various values of λ and for $\sigma^2 = 1$. As in Table 2, one interpolates over λ, whereas for $\sigma^2 \neq 1$ the quantiles are multiplied by σ.

Table 7 gives our results for the size of the test (proportion of rejections when the null is true) for the case that $\lambda = 1$ ($\theta = \sigma_v$). We considered only one value of λ because the relative variances of noise and inefficiency did not have much effect on the size of the test in Section 6.1. Table 7 has the same format as Table 3C did for the normal / half-normal case. The conclusions are also very similar to those for the normal / half normal case. The tests that use bootstrapped critical values are all quite accurate. The Tauchen version of the Pearson test is less reliable, and shows noticeable size distortions unless the sample size is very large or the number of cells is small.

Table 8 gives the power of the test of the exponential null against half normal alternatives. Specifically, in these cases v is N(0,1) and u is half normal with parameter (pre-truncation variance) σ_u^2. This implies that the mean of u is $\sqrt{2/\pi}\sigma_u$ = $0.798\sigma_u$ and the variance of u is $\frac{\pi-2}{\pi}\sigma_u^2$ = $0.363\sigma_u^2$ (so that the standard deviation of u is $0.603\sigma_u$).

Table 8 is the converse of Table 4A, in which the null was normal / half-normal and the alternative was normal / exponential. The results are similar to those in Table 4A. Power increases as n increases. Power also increases as σ_u increases, that is, as the importance of noise relative to inefficiency diminishes. The KS test using bootstrapped critical values is generally the most powerful test.

It appears that power is generally higher in Table 4A than in Table 8. That is, the test of the null of half-normal inefficiency against the alternative of exponential is more powerful than the test of the null of exponential inefficiency against the alternative of half-normal. However, this conclusion requires some care in matching cases in the two tables. In Table 4A, inefficiency (u) is exponential and the tabulation is in terms of θ, which is both the mean and the standard deviation of u. In Table 8, u is half normal and the tabulation is in terms of σ_u, but the mean of u is $0.798\sigma_u$ and the standard deviation is $0.603\sigma_u$. So, for example, $\theta = 2$ in Table 4A corresponds to σ_u = 2.506 if we match the expected value of u across the two tables, or it corresponds to σ_u = 3.316 if we match the standard deviation of u across the two tables. However, even taking that into account, the conclusion above still stands. The test of the null of half-normal inefficiency against the alternative of exponential is more powerful than the test of the null of exponential inefficiency against the alternative of half-normal.

Tables 9A and 9B give the power of the test of the exponential null against gamma alternatives. Specifically, in these cases v is N(0,1) and u is distributed as c times a gamma distribution with parameter m = 0.5 and 2, respectively. These two tables can be compared to Tables 5B and 5C, in which the null was that inefficiency was half-normal.

Some of the conclusions are as expected. The bootstrapped version of the KS test is generally most powerful. Power grows with n and with c. Interestingly, power is higher in Table 5B than in Table 9A, but power is higher in Table 9B than in Table 5C. That is, if inefficiency is distributed as gamma with $m = 0.5$, it is easier to reject the half-normal null than the exponential null. However, if inefficiency is distributed as gamma with $m = 2$, the power comparison is reversed. Presumably this is because a gamma distribution with $m = 0.5$ is closer to an exponential distribution than it is to a half-normal, and vice-versa for a gamma distribution with $m = 2$.

7. EMPIRICAL EXAMPLES

In this section we present two empirical examples.

Our first example is from Greene (2008), pp. 541-2. The data consist of $n = 25$ statewide observations on output and two inputs for the transportation equipment manufacturing industry, and were previously used by Zellner and Revankar (1970). The data are available online as described in Greene, Table F14.1, p. 1088. Greene, p. 542, gives the MLE for the normal/half-normal model and for the normal/exponential model. In both cases the variance of the one-sided error is about one-third of the total variance of the composed error, so there is a fair amount of statistical noise. The choice of distribution matters moderately here. For example, the state of Michigan ranks 15^{th} with an inefficiency of 0.1581 using the half-normal model, while it ranks 13^{th} with an inefficiency of 0.1076 using the exponential model, and these differences across models are similar for other states.

In this example we fail to reject the half-normal null. The Tauchen Pearson χ^2 statistic based on three cells equals 0.537, which is less than the usual critical value of 5.99 for the χ^2 distribution with two degrees of freedom, and also less than the bootstrapped critical value of 6.58.

The Kolmogorov-Smirnov statistic equals 0.0985, which is less than the bootstrapped critical value of 0.165. We also fail to reject the exponential null. The χ^2 statistic of 1.27 is less than 5.99 or the bootstrapped critical value of 6.92, and the Kolmogorov-Smirnov statistic of 0.0960 is less than the bootstrapped critical value of 0.166. These failures to reject are not surprising given the small sample size and the small variance of the one-sided error relative to the variance of noise.

Our second example is taken from Coelli et al. (2005), Chapter 9. The data are n = 344 observations on Philippine rice producers. The data are actually a panel for 43 producers over 8 years, but we follow Coelli et al. in treating them as a single cross-section. That is, as they did, we make the unrealistic assumption that errors are independent over time for a given producer, as well as across producers. The production function is translog with three inputs. Further details on the data can be found in Coelli et al., Appendix 2. The data are available online as described in the Preface (p. xvi) of Coelli et al.

Coelli et al., Sections 9.3-9.4, give the MLE for the normal/half-normal and normal/exponential models, as well as the model in which technical inefficiency is general truncated normal. The choice of distributional assumption matters a little more in this example than in the previous one. For example, for the first observation, the estimate of the one-sided error is 0.2635 under the half-normal assumption and 0.1744 under the exponential assumption. In this example, technical efficiency is much larger relative to noise than in the previous example. The variance of the one-sided error is 73% (half-normal model) or 65% (exponential model) of the total variance of the composed error, as compared to about one third in the previous example.

In this example, we fail to reject the half-normal null using the Pearson (Tauchen) test. The statistic based on three cells equals 3.99, which is less than the 5% critical value of 5.99. However, we do reject the half-normal null using the Kolmogorov-Smirnov test. The statistic is 0.0566, which is larger than the bootstrapped 5% critical value of 0.0506. So there is some evidence against the normal-half normal model. We fail to reject the exponential null at the 5% level using either test (compare 3.56 to 5.99, and 0.0391 to 0.0460). So the tests do not provide evidence against the normal-exponential model.

8. CONCLUDING REMARKS

In this paper we have considered goodness of fit tests for the stochastic frontier model. We are interested in testing the distributional assumption for the one-sided error (inefficiency term). The essential difficulty is that we can only observe the composed error, which is the sum of the one-sided error and normal random noise. So in the end we test the hypothesis that the composed error has the distribution that is implied by normality of the noise and the assumed distribution for the one-sided error.

We considered Pearson χ^2 goodness of fit tests based on expected and actual numbers of observations in cells defined by values of the composed error, and also the Kolmogorov-Smirnov test. We discussed the asymptotic theory that corrects the Pearson test for the effects of parameter estimation. We also noted that asymptotically correct critical values can be found by boostrapping, for either the Pearson test or the Kolmogorov-Smirnov test.

We performed simulations to investigate the size and power properties of the tests. In terms of size, bootstrapping works better than asymptotic theory. In terms of power, the Kolmogorov-Smirnov test dominates the Pearson tests, so that the best test overall appears to be the Kolmogorov-Smirnov test using critical values from the bootstrap.

The remaining problem is that the power of these tests against plausible alternative distributions is somewhat low. Reasonable power seems to require sample sizes and/or signal to noise ratios that are not commonly found in empirical applications. Making the same point somewhat differently, it is easy to distinguish an exponential distribution from a half-normal. However, it is hard to distinguish the sum of a normal and an exponential from the sum of a normal and a half-normal, unless the variance of the normal component is very small or the sample size is very large.

Further research is needed to understand the empirical significance of these findings. Philosophically, it does not matter if different models yield different results if we can distinguish statistically between the models; conversely, it does not matter if we cannot distinguish statistically between models, if the models give more or less the same results. It is only a problem if we cannot distinguish statistically between models *and* the models give substantively different results. Intuitively, it seems reasonable to conjecture that data sets for which it is hard to distinguish between different distributions of inefficiency are also data sets for which the different distributions lead to similar empirical results. (Presumably these are cases in which different distributions for inefficiency lead to essentially the same distribution of the composed error ε.) Therefore the relationship between robustness of results and the power of goodness of fit tests (or, more generally, the ability of any model selection method to distinguish between different distributions of inefficiency) is obviously an important issue to investigate.

APPENDIX A

In this Appendix we establish equation (8) of the text. We write $\bar{g}(\theta_0) = P - \hat{P}$, where P is the (k-1)-dimensional vector with j^{th} element $p_j = p_j(\theta_0)$ and $\hat{P}$ is the (k-1)-dimensional vector with j^{th} element $\hat{p}_j = O_j / n$. Also we write $V(\theta_0) = \Pi - PP'$ where Π is the diagonal matrix with j^{th} diagonal element equal to p_j. Now we use the fact (e.g. Abadir and Magnus (2005), p. 87) that

(A1) $$[\Pi - PP']^{-1} = \Pi^{-1} + \frac{1}{1 - P'\Pi^{-1}P}\Pi^{-1}PP'\Pi^{-1}$$

Therefore

(A2) $$n\bar{g}(\theta_0)'V(\theta_0)^{-1}\bar{g}(\theta_0) = n(\hat{P} - P)'\Pi^{-1}(\hat{P} - P) + \frac{n}{1 - P'\Pi^{-1}P}(\hat{P} - P)'\Pi^{-1}PP'\Pi^{-1}(\hat{P} - P)$$

The first term on the right hand side of (A2) equals $n\sum_{j=1}^{k-1}(\hat{p}_j - p_j)^2 / p_j = \sum_{j=1}^{k-1}(O_j - E_j)^2 / E_j$. For the second term, note that $1 - P'\Pi^{-1}P = 1 - \sum_{j=1}^{k-1} p_j = p_k$ and that $(\hat{P} - P)'\Pi^{-1}P =$ $(\hat{P} - P)'e_{k-1}$ (where e_{k-1} is a vector of dimension (k-1) with each element equal to one) = $[(1 - \hat{p}_k) - (1 - p_k)] = (p_k - \hat{p}_k)$. Therefore $n\bar{g}(\theta_0)'V(\theta_0)^{-1}\bar{g}(\theta_0) = \sum_{j=1}^{k-1}(O_j - E_j)^2 / E_j$ + $n(p_k - \hat{p}_k)^2 / p_k = \sum_{j=1}^{k}(O_j - E_j)^2 / E_j$.

APPENDIX B

In this Appendix we discuss the goodness of fit test based on quantiles and its relationship to the Pearson test based on actual and expected cell counts. Suppose that we pick (k-1) probabilities $0 < p_1 < p_2 \cdots < p_{k-1} < 1$. Let the corresponding population quantiles be $m_1(\theta) < m_2(\theta) \cdots < m_{k-1}(\theta)$, so that $P(y \le m_j(\theta)) = p_j$, and let the sample quantiles be

$\hat{m}_1 \le \hat{m}_2 \cdots \le \hat{m}_{k-1}$. So now the test will depend on ($\hat{m} - m$), the vector whose j^{th} element equals ($\hat{m}_j - m_j(\theta)$), and the test statistic equals $n(\hat{m} - m(\hat{\theta}))' W(\hat{m} - m(\hat{\theta}))$ with an appropriate choice of W.

To see how this compares to the CMT test, we note that $\sqrt{n}(\hat{m}_j - m_j(\theta))$ is asymptotically normal, and so it must be expressable as an average (plus an asymptotically negligible term). This is the "influence function representation," which is given by:

(B1) $$\sqrt{n}(\hat{m}_j - m_j(\theta)) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} r_{ij}(\theta) + o_p(1)$$

where $o_p(1)$ is an asymptotically negligible term (i.e., it converges in probability to zero), and where

(B2) $$r_{ij}(\theta) = \frac{1}{f(m_j(\theta))} [p_j - 1(y_i \le m_j(\theta))]$$

where f is the pdf of y. See, for example, Ruppert and Carroll (1980), p. 832. Therefore the test based on ($\hat{m} - m$) is equivalent in large samples to the CMT test based on the moment conditions $E[1(y \le m_j(\theta)) - p_j]$, j = 1, 2, … k-1. This is an overlapping set of cells. However, it is also equivalent to consider the non-overlapping cells: $A_1 = \{y | y \le m_1(\theta)\}$, $A_2 = \{y | m_1(\theta) < y \le m_2(\theta)\}$, etc. The resulting test is the CMT test based on observed versus actual cell counts, as discussed in the text.

APPENDIX C

In this Appendix we derive analytically the variance matrix C used in the conditional moment test, for the case of a normal distribution. We wish to evaluate

(C1) $$C_{11} = E(ss') \;\;,\;\; C_{12} = E(sg') \;\;,\;\; C_{22} = E(gg') \;.$$

Here $s = s(y,\theta)$ is the score function for the normal distribution, given by

$$\text{(C2)} \qquad s(y,\theta) = \begin{bmatrix} \dfrac{1}{\sigma^2}(y-\mu) \\ \dfrac{-1}{2\sigma^2} + \dfrac{1}{2\sigma^4}(y-\mu)^2 \end{bmatrix}$$

and $g = g(y,\theta)$ is the vector whose j^{th} element equals $[1(y \in A_j) - p_j]$.

It is well known that C_{11} is the information matrix for the normal distribution, given by

$$\text{(C3)} \qquad \begin{bmatrix} \dfrac{1}{\sigma^2} & 0 \\ 0 & \dfrac{1}{2\sigma^4} \end{bmatrix}.$$

Also C_{22} equals the matrix $V(\theta)$ as defined in the discussion following equation (6) of the text.

This leaves the submatrix C_{12}. It is of dimension 2 by (*k*-1). We will evaluate in turn the (1,*j*) and (2,*j*) elements of this matrix. To do so we make the reasonable assumption that the cells are intervals, so that $A_j = (a,b]$, where for notational simplicity we do not express the subscript "*j*" that should appear on *a* and *b*.

Then element (1,*j*) of C_{12} equals

$$\frac{1}{\sigma^2}E(y-\mu)[1(y \in A_j) - p_j] = \frac{1}{\sigma^2}Ey[1(y \in A_j) - p_j]$$

$$= \frac{1}{\sigma^2}Ey1(y \in A_j) - \frac{1}{\sigma^2}p_j\mu$$

$$= \frac{p_j}{\sigma^2}[E(y|a < y \le b) - \mu]$$

$$= \frac{1}{\sigma^2}\left[\phi\left(\frac{a-\mu}{\sigma}\right) - \phi\left(\frac{b-\mu}{\sigma}\right)\right],$$

where "ϕ" is the standard normal density function. Here we have evaluated the conditioanal expectation $E(y|a<y\le b)=\mu+\frac{1}{p_j}\left[\phi\left(\frac{a-\mu}{\sigma}\right)-\phi\left(\frac{b-\mu}{\sigma}\right)\right]$ as in Johnson and Kotz (1970), equation (79), p. 81.

Similarly element (2,*j*) of C_{12} equals

$$E\left[\frac{-1}{2\sigma^2}+\frac{1}{2\sigma^4}(y-\mu)^2\right][1(y\in A_j)-p_j]$$

$$=E\left[\frac{1}{2\sigma^4}(y-\mu)^2\right][1(y\in A_j)-p_j]$$

$$=\frac{1}{2\sigma^4}E(y-\mu)^2 1(a<y\le b)-\frac{p_j}{2\sigma^4}$$

$$=\frac{1}{2\sigma^4}Ey^2 1(y\in A_j)+\frac{1}{2\sigma^4}(-2\mu)Ey1(y\in A_j)+\frac{1}{2\sigma^4}\mu^2 p_j-\frac{p_j}{2\sigma^4}$$

$$=\frac{1}{2\sigma^4}Ey^2 1(y\in A_j)-\frac{\mu}{\sigma^4}p_j E(y|y\in A_j)+\frac{p_j\mu^2}{2\sigma^4}-\frac{p_j}{2\sigma^4},$$

where $Ey^2 1(y\in A_j)=p_j\,\mathrm{var}(y|y\in A_j)+p_j\left[E(y|y\in A_j)\right]^2$. Furthermore,

$$Ey^2 1(y\in A_j)=p_j\sigma^2\left\{1-\frac{b\phi(b)-a\phi(a)}{\Phi(b)-\Phi(a)}-\left[\frac{\phi(b)-\phi(a)}{\Phi(b)-\Phi(a)}\right]^2\right\}+p_j\left[\mu-\sigma\frac{\phi(b)-\phi(a)}{\Phi(b)-\Phi(a)}\right]^2.$$

TABLE 1A

Size of the test of the hypothesis that the data are N(0,1)

Nominal size = 0.05

k	*n*	Pearson	Bootstrap Pearson	KS	Bootstrap KS
3	50	0.049	0.053	0.040	0.045
	100	0.054	0.047	0.040	0.048
	250	0.053	0.039	0.046	0.048
	500	0.047	0.045	0.050	0.056
5	50	0.043	0.050	*	*
	100	0.049	0.044	*	*
	250	0.049	0.041	*	*
	500	0.051	0.058	*	*
10	50	0.048	0.052	*	*
	100	0.049	0.043	*	*
	250	0.052	0.056	*	*
	500	0.051	0.058	*	*

* The number of cells (k) is not relevant for the Kolmogorov-Smirnov test.

TABLE 1B

Size of the test of the hypothesis that the data are normal

Nominal size = 0.05

k	*n*	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS	Simple Hypothesis OPG	Composite Hypothesis *C* Known
3	50	0.067	0.056	0.051	0.057	0.050	0.049
	100	0.055	0.052	0.053	0.054	0.050	0.050
	250	0.052	0.050	0.049	0.047	0.054	0.045
	500	0.053	0.058	0.053	0.048	0.048	0.051
5	50	0.086	0.055	0.044	*	0.067	0.049
	100	0.070	0.042	0.055	*	0.059	0.050
	250	0.057	0.043	0.060	*	0.053	0.050
	500	0.061	0.054	0.062	*	0.053	0.048
10	50	0.185	0.054	0.060	*	0.139	0.044
	100	0.117	0.053	0.050	*	0.101	0.049
	250	0.079	0.052	0.042	*	0.071	0.051
	500	0.060	0.044	0.058	*	0.059	0.047

* The number of cells (k) is not relevant for the Kolmogorov-Smirnov test.

TABLE 2

Quantiles of the distribution of the normal / half normal composed error

$$\boldsymbol{\sigma^2 = \sigma_u^2 + \sigma_v^2 = 1}\text{ , various }\boldsymbol{\lambda = \sigma_u/\sigma_v}$$

λ	Quantile .10	.20	.30	.40	.50	.60	.70	.80	.90
0.0	-1.281	-0.841	-0.524	-0.253	0.000	0.253	0.524	0.841	1.281
0.1	-1.357	-0.918	-0.602	-0.332	-0.080	0.173	0.444	0.759	1.198
0.2	-1.423	-0.987	-0.675	-0.407	-0.156	0.094	0.362	0.675	1.109
0.3	-1.477	-1.048	-0.739	-0.475	-0.228	0.018	0.282	0.590	1.017
0.4	-1.522	-1.099	-0.796	-0.537	-0.294	-0.053	0.206	0.508	0.926
0.5	-1.556	-1.141	-0.843	-0.589	-0.353	-0.117	0.135	0.430	0.837
0.6	-1.582	-1.176	-0.884	-0.635	-0.405	-0.174	0.071	0.358	0.754
0.7	-1.602	-1.202	-0.917	-0.674	-0.449	-0.224	0.014	0.292	0.675
0.8	-1.616	-1.223	-0.943	-0.706	-0.486	-0.269	-0.037	0.233	0.603
0.9	-1.626	-1.238	-0.964	-0.733	-0.518	-0.306	-0.081	0.180	0.537
1.0	-1.632	-1.250	-0.981	-0.754	-0.545	-0.339	-0.120	0.133	0.478
1.1	-1.637	-1.259	-0.994	-0.772	-0.567	-0.366	-0.153	0.091	0.425
1.2	-1.640	-1.266	-1.004	-0.786	-0.586	-0.389	-0.183	0.054	0.377
1.3	-1.642	-1.271	-1.012	-0.797	-0.601	-0.409	-0.209	0.022	0.334
1.4	-1.643	-1.274	-1.018	-0.806	-0.614	-0.426	-0.230	-0.006	0.295
1.5	-1.644	-1.276	-1.023	-0.814	-0.625	-0.440	-0.249	-0.032	0.261
1.6	-1.644	-1.278	-1.026	-0.819	-0.633	-0.453	-0.266	-0.054	0.230
1.7	-1.644	-1.279	-1.029	-0.824	-0.640	-0.463	-0.280	-0.074	0.202
1.8	-1.645	-1.280	-1.031	-0.828	-0.646	-0.472	-0.293	-0.091	0.177
1.9	-1.645	-1.280	-1.033	-0.831	-0.651	-0.480	-0.304	-0.107	0.154
2.0	-1.645	-1.281	-1.034	-0.833	-0.656	-0.486	-0.313	-0.120	0.134
2.1	-1.645	-1.281	-1.034	-0.835	-0.659	-0.491	-0.322	-0.132	0.115
2.2	-1.645	-1.281	-1.035	-0.837	-0.662	-0.496	-0.329	-0.144	0.098
2.3	-1.645	-1.282	-1.035	-0.838	-0.664	-0.500	-0.335	-0.154	0.083
2.4	-1.645	-1.282	-1.036	-0.838	-0.666	-0.504	-0.341	-0.162	0.069
2.5	-1.645	-1.282	-1.036	-0.839	-0.667	-0.507	-0.346	-0.170	0.056
2.6	-1.645	-1.282	-1.036	-0.840	-0.669	-0.509	-0.350	-0.177	0.044
2.7	-1.645	-1.282	-1.036	-0.840	-0.670	-0.512	-0.354	-0.184	0.034
2.8	-1.645	-1.282	-1.036	-0.841	-0.671	-0.513	-0.358	-0.190	0.024
2.9	-1.645	-1.282	-1.036	-0.841	-0.672	-0.515	-0.361	-0.195	0.014
3.0	-1.645	-1.282	-1.036	-0.841	-0.672	-0.516	-0.364	-0.200	0.006

λ	Quantile								
	.10	.20	.30	.40	.50	.60	.70	.80	.90
3.1	-1.645	-1.282	-1.036	-0.841	-0.673	-0.518	-0.366	-0.204	-0.002
3.2	-1.645	-1.282	-1.036	-0.841	-0.673	-0.519	-0.368	-0.208	-0.009
3.3	-1.645	-1.282	-1.036	-0.841	-0.673	-0.519	-0.370	-0.212	-0.016
3.4	-1.645	-1.282	-1.036	-0.841	-0.674	-0.520	-0.372	-0.215	-0.022
3.5	-1.645	-1.282	-1.036	-0.841	-0.674	-0.521	-0.373	-0.218	-0.027
3.6	-1.645	-1.282	-1.036	-0.841	-0.674	-0.521	-0.374	-0.221	-0.033
3.7	-1.645	-1.282	-1.036	-0.841	-0.674	-0.522	-0.376	-0.224	-0.038
3.8	-1.645	-1.282	-1.036	-0.841	-0.674	-0.522	-0.377	-0.226	-0.043
3.9	-1.645	-1.282	-1.036	-0.841	-0.674	-0.523	-0.378	-0.228	-0.047
4.0	-1.645	-1.282	-1.036	-0.842	-0.674	-0.523	-0.379	-0.230	-0.051
4.1	-1.645	-1.282	-1.036	-0.842	-0.674	-0.523	-0.379	-0.232	-0.055
4.2	-1.645	-1.282	-1.036	-0.842	-0.674	-0.523	-0.380	-0.233	-0.059
4.3	-1.645	-1.282	-1.036	-0.842	-0.674	-0.523	-0.381	-0.235	-0.062
4.4	-1.645	-1.282	-1.036	-0.842	-0.674	-0.524	-0.381	-0.236	-0.065
4.5	-1.645	-1.282	-1.036	-0.842	-0.674	-0.524	-0.382	-0.238	-0.068
4.6	-1.645	-1.282	-1.036	-0.842	-0.674	-0.524	-0.382	-0.239	-0.071
4.7	-1.645	-1.282	-1.036	-0.842	-0.674	-0.524	-0.383	-0.240	-0.073
4.8	-1.645	-1.282	-1.036	-0.842	-0.674	-0.524	-0.383	-0.241	-0.076
4.9	-1.645	-1.282	-1.036	-0.842	-0.674	-0.524	-0.383	-0.242	-0.078
5.0	-1.645	-1.282	-1.036	-0.842	-0.674	-0.524	-0.383	-0.243	-0.081
5.1	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.383	-0.244	-0.083
5.2	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.384	-0.244	-0.085
5.3	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.384	-0.245	-0.086
5.4	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.384	-0.246	-0.088
5.5	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.384	-0.246	-0.090
5.6	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.384	-0.247	-0.092
5.7	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.384	-0.247	-0.093
5.8	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.248	-0.094
5.9	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.248	-0.096
6.0	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.248	-0.097
6.1	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.249	-0.098
6.2	-1.645	-1.282	-1.036	-0.842	-0.675	-0.525	-0.385	-0.249	-0.100
6.3	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.250	-0.101
6.4	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.250	-0.102
6.5	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.250	-0.103
6.6	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.250	-0.104
6.7	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.251	-0.105
6.8	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.251	-0.105
6.9	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.251	-0.106
7.0	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.251	-0.107

λ	Quantile								
	.10	.20	.30	.40	.50	.60	.70	.80	.90
7.1	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.251	-0.108
7.2	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.251	-0.109
7.3	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.252	-0.109
7.4	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.252	-0.110
7.5	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.252	-0.111
7.6	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.252	-0.111
7.7	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.252	-0.112
7.8	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.252	-0.112
7.9	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.252	-0.113
8.0	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.252	-0.113
8.1	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.252	-0.114
8.2	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.114
8.3	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.115
8.4	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.115
8.5	-1.645	-1.282	-1.036	-0.842	-0.675	-0.525	-0.385	-0.253	-0.116
8.6	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.116
8.7	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.116
8.8	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.117
8.9	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.117
9.0	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.117
9.1	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.118
9.2	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.118
9.3	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.118
9.4	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.119
9.5	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.119
9.6	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.119
9.7	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.119
9.8	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.120
9.9	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.120
10	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.120
11	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.122
12	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.123
13	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.124
14	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.124
15	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.125
20	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.125
50	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.125
100	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.125
1000	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.126
10000	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.126

TABLE 3A

Size of the test of the hypothesis that the data are normal / half-normal

$\boldsymbol{\lambda = \sigma_u/\sigma_v = 0.1}$

Nominal size = 0.05

k	*n*	Wrong Skew (%)	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
3	50	50.2	0.077	0.051	0.057	0.048
	100	50.9	0.062	0.053	0.055	0.062
	250	50.2	0.056	0.049	0.040	0.041
5	50	50.2	0.092	0.043	0.043	*
	100	50.9	0.068	0.046	0.046	*
	250	50.2	0.060	0.049	0.049	*
10	50	50.2	0.200	0.045	0.047	*
	100	50.9	0.119	0.049	0.054	*
	250	50.2	0.076	0.051	0.044	*

*The number of cells (k) is not relevant for the Kolmogorov-Smirnov test.

TABLE 3B

Size of the test of the hypothesis that the data are normal / half-normal

λ = 0.5

Nominal size = 0.05

k	*n*	Wrong Skew (%)	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
3	50	47.8	0.079	0.048	0.031	0.047
	100	47.4	0.055	0.037	0.041	0.050
	250	44.3	0.051	0.050	0.046	0.045
5	50	47.8	0.089	0.044	0.044	*
	100	47.4	0.065	0.056	0.055	*
	250	44.3	0.055	0.044	0.057	*
10	50	47.8	0.191	0.053	0.053	*
	100	47.4	0.113	0.056	0.055	*
	250	44.3	0.073	0.055	0.049	*

* The number of cells (k) is not relevant for the Kolmogorov-Smirnov test.

TABLE 3C

Size of the test of the hypothesis that the data are normal / half-normal

$\lambda = 1$

Nominal size = 0.05

k	*n*	Wrong Skew (%)	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
3	50	36.0	0.080	0.051	0.041	0.042
	100	30.8	0.059	0.053	0.037	0.041
	250	19.5	0.049	0.048	0.053	0.040
5	50	36.0	0.093	0.042	0.039	*
	100	30.8	0.065	0.056	0.042	*
	250	19.5	0.053	0.043	0.043	*
10	50	36.0	0.190	0.039	0.050	*
	100	30.8	0.111	0.041	0.044	*
	250	19.5	0.068	0.061	0.055	*

* The number of cells (k) is not relevant for the Kolmogorov-Smirnov test.

TABLE 3D

Size of the test of the hypothesis that the data are normal / half-normal

$\lambda = 2$

Nominal size = 0.05

k	*n*	Wrong Skew (%)	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
3	50	11.5	0.074	0.053	0.034	0.037
	100	4.1	0.067	0.046	0.039	0.042
	250	0.2	0.050	0.056	0.044	0.045
5	50	11.5	0.107	0.043	0.042	*
	100	4.1	0.072	0.064	0.048	*
	250	0.2	0.055	0.060	0.053	*
10	50	11.5	0.233	0.040	0.043	*
	100	4.1	0.122	0.048	0.046	*
	250	0.2	0.069	0.058	0.052	*

* The number of cells (k) is not relevant for the Kolmogorov-Smirnov test.

TABLE 3E

Size of the test of the hypothesis that the data are normal / half-normal

$\lambda = 10$

Nominal size = 0.05

k	*n*	Wrong Skew (%)	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
3	50	0.1	0.060	0.049	0.050	0.044
	100	0	0.053	0.057	0.051	0.038
	250	0	0.051	0.048	0.048	0.045
5	50	0.1	0.090	0.041	0.051	*
	100	0	0.064	0.049	0.048	*
	250	0	0.059	0.053	0.048	*
10	50	0.1	0.238	0.038	0.045	*
	100	0	0.116	0.054	0.057	*
	250	0	0.073	0.052	0.043	*

* The number of cells (k) is not relevant for the Kolmogorov-Smirnov test.

TABLE 4A

Power of the test of the hypothesis that the data are normal / half-normal

Alternative: the data are normal / exponential (mean = θ)

Nominal size = 0.05

Number of cells: k = 3

n	θ	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
50	0.1	0.069	0.059	0.059	0.046
	0.5	0.071	0.055	0.053	0.041
	1	0.079	0.056	0.059	0.054
	2	0.142	0.110	0.098	0.147
	5	0.269	0.217	0.195	0.340
	10	0.367	0.285	0.294	0.494
100	0.1	0.064	0.049	0.054	0.056
	0.5	0.060	0.055	0.050	0.048
	1	0.080	0.084	0.069	0.089
	2	0.196	0.188	0.207	0.239
	5	0.468	0.440	0.384	0.530
	10	0.633	0.563	0.537	0.700
250	0.1	0.064	0.049	0.051	0.057
	0.5	0.052	0.051	0.041	0.048
	1	0.101	0.084	0.108	0.150
	2	0.416	0.407	0.476	0.507
	5	0.879	0.842	0.789	0.930
	10	0.959	0.945	0.922	0.966

TABLE 4B

Power of the test of the hypothesis that the data are normal / half-normal

Alternative: the data are normal / exponential (mean = θ)

Nominal size = 0.05

Number of cells: k = 5

n	θ	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
50	0.1	0.086	0.058	0.052	0.046
	0.5	0.087	0.042	0.057	0.041
	1	0.097	0.065	0.064	0.054
	2	0.150	0.097	0.079	0.147
	5	0.265	0.137	0.123	0.340
	10	0.343	0.167	0.185	0.494
100	0.1	0.062	0.054	0.055	0.056
	0.5	0.087	0.045	0.044	0.048
	1	0.107	0.076	0.071	0.089
	2	0.202	0.167	0.135	0.239
	5	0.399	0.354	0.282	0.530
	10	0.537	0.435	0.383	0.700
250	0.1	0.061	0.056	0.051	0.057
	0.5	0.059	0.061	0.056	0.048
	1	0.110	0.101	0.091	0.150
	2	0.383	0.382	0.302	0.507
	5	0.851	0.799	0.695	0.930
	10	0.944	0.917	0.833	0.966

TABLE 4C

Power of the test of the hypothesis that the data are normal / half-normal

Alternative: the data are normal / exponential (mean = θ)

Nominal size = 0.05

Number of cells: k = 10

n	θ	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
50	0.1	0.191	0.044	0.047	0.046
	0.5	0.167	0.049	0.045	0.041
	1	0.203	0.050	0.037	0.054
	2	0.245	0.065	0.063	0.147
	5	0.370	0.079	0.102	0.340
	10	0.485	0.093	0.110	0.494
100	0.1	0.119	0.050	0.042	0.056
	0.5	0.104	0.049	0.047	0.048
	1	0.140	0.061	0.061	0.089
	2	0.225	0.108	0.101	0.239
	5	0.364	0.209	0.210	0.530
	10	0.499	0.295	0.260	0.700
250	0.1	0.082	0.052	0.045	0.057
	0.5	0.077	0.045	0.052	0.048
	1	0.124	0.075	0.059	0.150
	2	0.338	0.267	0.217	0.507
	5	0.716	0.671	0.494	0.930
	10	0.879	0.815	0.789	0.966

TABLE 5A

Power of the test of the hypothesis that the data are normal / half-normal

Alternative: the data are normal / gamma (*u* is *c* times gamma(*m*))

Nominal size = 0.05

***m* = 0.1**

n	*c*	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
50	0.1	0.069	0.051	0.046	0.042
	0.5	0.060	0.061	0.061	0.051
	1	0.073	0.065	0.066	0.058
	2	0.081	0.135	0.120	0.117
	5	0.407	0.365	0.403	0.467
	10	0.785	0.695	0.730	0.814
100	0.1	0.059	0.055	0.046	0.055
	0.5	0.061	0.052	0.049	0.068
	1	0.075	0.046	0.046	0.056
	2	0.099	0.095	0.127	0.177
	5	0.614	0.602	0.607	0.978
	10	0.962	0.960	0.965	0.984
250	0.1	0.062	0.058	0.060	0.060
	0.5	0.052	0.055	0.051	0.052
	1	0.063	0.060	0.062	0.059
	2	0129	0.134	0.175	0.308
	5	0.888	0.921	0.954	1.000
	10	1.000	1.000	1.000	1.000

TABLE 5B

Power of the test of the hypothesis that the data are normal / half-normal

Alternative: the data are normal / gamma (*u* is *c* times gamma(*m*))

Nominal size = 0.05

***m* = 0.5**

n	*c*	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
50	0.1	0.068	0.054	0.049	0.061
	0.5	0.049	0.046	0.046	0.046
	1	0.073	0.066	0.066	0.056
	2	0.153	0.132	0.132	0.154
	5	0.511	0.416	0.416	0.621
	10	0.791	0.740	0.740	0.887
100	0.1	0.074	0.061	0.061	0.053
	0.5	0.055	0.061	0.061	0.050
	1	0.061	0.078	0.078	0.078
	2	0.260	0.271	0.271	0.328
	5	0.784	0.732	0.732	0.869
	10	0.974	0.948	0.948	0.945
250	0.1	0.067	0.053	0.053	0.060
	0.5	0.061	0.069	0.069	0.067
	1	0.080	0.082	0.082	0.120
	2	0.516	0.583	0.583	0.685
	5	0.995	0.973	0.973	1.000
	10	1.000	1.000	1.000	1.000

TABLE 5C

Power of the test of the hypothesis that the data are normal / half-normal

Alternative: the data are normal / gamma (*u* is *c* times gamma(*m*))

Nominal size = 0.05

***m* = 2**

n	*c*	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
50	0.1	0.058	0.054	0.061	0.043
	0.5	0.050	0.038	0.043	0.029
	1	0.073	0.057	0.053	0.050
	2	0.072	0.075	0.067	0.049
	5	0.095	0.048	0.064	0.073
	10	0.110	0.076	0.070	0.082
100	0.1	0.071	0.053	0.050	0.060
	0.5	0.044	0.063	0.049	0.046
	1	0.066	0.062	0.053	0.068
	2	0.091	0.075	0.092	0.107
	5	0.106	0.096	0.094	0.127
	10	0.125	0.110	0.107	0.132
250	0.1	0.040	0.061	0.057	0.057
	0.5	0.049	0.052	0.054	0.063
	1	0.080	0.088	0.079	0.100
	2	0.164	0.141	0.158	0.198
	5	0.210	0.196	0.188	0.220
	10	0.219	0.210	0.217	0.233

TABLE 5D

Power of the test of the hypothesis that the data are normal / half-normal

Alternative: the data are normal / gamma (*u* is *c* times gamma(*m*))

Nominal size = 0.05

***m* = 10**

n	*c*	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
50	0.1	0.061	0.050	0.052	0.060
	0.5	0.062	0.054	0.043	0.055
	1	0.065	0.042	0.052	0.072
	2	0.065	0.054	0.044	0.068
	5	0.095	0.074	0.076	0.098
	10	0.110	0.098	0.099	0.115
100	0.1	0.062	0.051	0.049	0.055
	0.5	0.064	0.060	0.055	0.053
	1	0.065	0.048	0.052	0.068
	2	0.068	0.062	0.063	0.075
	5	0.093	0.070	0.065	0.099
	10	0.100	0.082	0.085	0.121
250	0.1	0.055	0.055	0.054	0.039
	0.5	0.059	0.063	0.059	0.055
	1	0.055	0.066	0.066	0.043
	2	0.071	0.069	0.072	0.058
	5	0.082	0.079	0.090	0.087
	10	0.104	0.095	0.088	0.119

TABLE 6

Quantiles of the distribution of the normal / exponential composed error

$$\boldsymbol{\sigma^2 = \theta^2 + \sigma_v^2 = 1}\text{, various } \boldsymbol{\lambda = \theta/\sigma_v}$$

λ	Quantile								
	.10	.20	.30	.40	.50	.60	.70	.80	.90
0.0	-1.281	-0.841	-0.524	-0.253	0.000	0.253	0.524	0.841	1.281
0.1	-1.381	-0.941	-0.624	-0.353	-0.099	0.154	0.425	0.742	1.182
0.2	-1.479	-1.036	-0.718	-0.447	-0.194	0.059	0.330	0.646	1.083
0.3	-1.570	-1.124	-0.805	-0.533	-0.280	-0.028	0.240	0.554	0.988
0.4	-1.654	-1.201	-0.880	-0.609	-0.357	-0.107	0.159	0.469	0.895
0.5	-1.728	-1.267	-0.944	-0.673	-0.423	-0.176	0.087	0.391	0.808
0.6	-1.793	-1.321	-0.997	-0.726	-0.479	-0.235	0.022	0.320	0.727
0.7	-1.848	-1.366	-1.039	-0.770	-0.525	-0.285	-0.033	0.257	0.653
0.8	-1.896	-1.402	-1.073	-0.805	-0.563	-0.328	-0.081	0.201	0.584
0.9	-1.938	-1.431	-1.100	-0.833	-0.594	-0.363	-0.123	0.152	0.523
1.0	-1.974	-1.454	-1.121	-0.855	-0.619	-0.393	-0.158	0.108	0.467
1.1	-2.006	-1.473	-1.137	-0.872	-0.639	-0.417	-0.188	0.070	0.417
1.2	-2.034	-1.489	-1.149	-0.886	-0.655	-0.437	-0.214	0.037	0.372
1.3	-2.059	-1.501	-1.159	-0.896	-0.668	-0.454	-0.236	0.008	0.332
1.4	-2.081	-1.512	-1.166	-0.903	-0.678	-0.467	-0.254	-0.017	0.296
1.5	-2.101	-1.521	-1.172	-0.909	-0.686	-0.479	-0.270	-0.040	0.264
1.6	-2.118	-1.529	-1.176	-0.914	-0.692	-0.488	-0.284	-0.059	0.235
1.7	-2.134	-1.535	-1.180	-0.917	-0.697	-0.496	-0.296	-0.077	0.209
1.8	-2.148	-1.541	-1.183	-0.919	-0.700	-0.502	-0.306	-0.092	0.185
1.9	-2.160	-1.546	-1.185	-0.921	-0.703	-0.507	-0.314	-0.106	0.164
2.0	-2.171	-1.551	-1.186	-0.921	-0.705	-0.511	-0.322	-0.118	0.144
2.1	-2.181	-1.555	-1.188	-0.921	-0.707	-0.514	-0.328	-0.128	0.127
2.2	-2.190	-1.559	-1.189	-0.921	-0.706	-0.517	-0.334	-0.138	0.111
2.3	-2.198	-1.562	-1.190	-0.921	-0.708	-0.519	-0.338	-0.146	0.097
2.4	-2.206	-1.566	-1.191	-0.921	-0.708	-0.521	-0.343	-0.154	0.083
2.5	-2.212	-1.569	-1.192	-0.921	-0.708	-0.522	-0.346	-0.161	0.071
2.6	-2.218	-1.571	-1.193	-0.921	-0.708	-0.523	-0.349	-0.167	0.060
2.7	-2.224	-1.574	-1.193	-0.921	-0.708	-0.524	-0.352	-0.173	0.050
2.8	-2.229	-1.576	-1.194	-0.922	-0.708	-0.525	-0.354	-0.178	0.040
2.9	-2.233	-1.578	-1.194	-0.922	-0.708	-0.525	-0.356	-0.182	0.032
3.0	-2.237	-1.580	-1.195	-0.922	-0.707	-0.525	-0.358	-0.186	0.024

λ	Quantile								
	.10	.20	.30	.40	.50	.60	.70	.80	.90
3.1	-2.241	-1.581	-1.195	-0.921	-0.707	-0.525	-0.359	-0.190	0.016
3.2	-2.244	-1.583	-1.196	-0.921	-0.706	-0.525	-0.360	-0.193	0.009
3.3	-2.247	-1.584	-1.196	-0.921	-0.706	-0.525	-0.361	-0.196	0.003
3.4	-2.250	-1.585	-1.196	-0.920	-0.706	-0.525	-0.362	-0.199	-0.003
3.5	-2.253	-1.587	-1.197	-0.920	-0.705	-0.525	-0.363	-0.202	-0.009
3.6	-2.255	-1.588	-1.197	-0.920	-0.704	-0.525	-0.364	-0.204	-0.014
3.7	-2.258	-1.589	-1.197	-0.920	-0.704	-0.524	-0.364	-0.206	-0.019
3.8	-2.260	-1.590	-1.198	-0.920	-0.703	-0.524	-0.365	-0.208	-0.023
3.9	-2.262	-1.591	-1.198	-0.919	-0.703	-0.524	-0.365	-0.210	-0.028
4.0	-2.264	-1.592	-1.198	-0.919	-0.702	-0.523	-0.365	-0.211	-0.031
4.1	-2.266	-1.592	-1.198	-0.919	-0.702	-0.523	-0.366	-0.213	-0.035
4.2	-2.267	-1.593	-1.199	-0.919	-0.701	-0.523	-0.366	-0.214	-0.039
4.3	-2.269	-1.594	-1.199	-0.919	-0.701	-0.522	-0.366	-0.215	-0.042
4.4	-2.271	-1.594	-1.199	-0.919	-0.701	-0.522	-0.366	-0.217	-0.045
4.5	-2.272	-1.595	-1.200	-0.919	-0.701	-0.522	-0.366	-0.217	-0.048
4.6	-2.273	-1.596	-1.200	-0.918	-0.700	-0.521	-0.366	-0.218	-0.051
4.7	-2.274	-1.596	-1.200	-0.918	-0.700	-0.521	-0.366	-0.219	-0.053
4.8	-2.275	-1.597	-1.200	-0.918	-0.700	-0.521	-0.366	-0.220	-0.056
4.9	-2.277	-1.597	-1.200	-0.918	-0.699	-0.520	-0.366	-0.221	-0.058
5.0	-2.278	-1.598	-1.200	-0.918	-0.699	-0.520	-0.366	-0.221	-0.060
5.1	-2.279	-1.598	-1.200	-0.918	-0.699	-0.520	-0.366	-0.222	-0.062
5.2	-2.279	-1.599	-1.200	-0.918	-0.699	-0.519	-0.366	-0.222	-0.064
5.3	-2.280	-1.599	-1.200	-0.918	-0.699	-0.519	-0.365	-0.223	-0.066
5.4	-2.281	-1.599	-1.201	-0.918	-0.698	-0.519	-0.365	-0.223	-0.068
5.5	-2.281	-1.600	-1.201	-0.918	-0.698	-0.519	-0.365	-0.224	-0.070
5.6	-2.282	-1.600	-1.201	-0.918	-0.698	-0.518	-0.365	-0.224	-0.071
5.7	-2.283	-1.600	-1.201	-0.918	-0.698	-0.518	-0.365	-0.224	-0.073
5.8	-2.283	-1.601	-1.201	-0.918	-0.698	-0.518	-0.365	-0.225	-0.074
5.9	-2.284	-1.601	-1.201	-0.918	-0.697	-0.518	-0.365	-0.225	-0.075
6.0	-2.285	-1.601	-1.201	-0.918	-0.697	-0.518	-0.364	-0.225	-0.077
6.1	-2.285	-1.601	-1.201	-0.917	-0.697	-0.517	-0.364	-0.225	-0.078
6.2	-2.286	-1.602	-1.201	-0.917	-0.697	-0.517	-0.364	-0.226	-0.079
6.3	-2.286	-1.602	-1.201	-0.917	-0.697	-0.517	-0.364	-0.226	-0.080
6.4	-2.287	-1.602	-1.201	-0.917	-0.697	-0.517	-0.364	-0.226	-0.081
6.5	-2.287	-1.602	-1.202	-0.917	-0.697	-0.517	-0.364	-0.226	-0.082
6.6	-2.288	-1.602	-1.202	-0.917	-0.697	-0.516	-0.363	-0.226	-0.083
6.7	-2.289	-1.603	-1.202	-0.917	-0.697	-0.516	-0.363	-0.226	-0.084
6.8	-2.289	-1.603	-1.202	-0.917	-0.697	-0.516	-0.363	-0.227	-0.085
6.9	-2.289	-1.603	-1.202	-0.917	-0.696	-0.516	-0.363	-0.227	-0.086
7.0	-2.289	-1.603	-1.202	-0.917	-0.696	-0.516	-0.363	-0.227	-0.087

λ	Quantile								
	.10	.20	.30	.40	.50	.60	.70	.80	.90
7.1	-2.290	-1.604	-1.202	-0.917	-0.696	-0.516	-0.363	-0.227	-0.088
7.2	-2.290	-1.604	-1.202	-0.917	-0.696	-0.516	-0.363	-0.227	-0.088
7.3	-2.290	-1.604	-1.202	-0.917	-0.696	-0.515	-0.363	-0.227	-0.089
7.4	-2.290	-1.604	-1.202	-0.917	-0.696	-0.515	-0.362	-0.227	-0.090
7.5	-2.291	-1.604	-1.202	-0.917	-0.696	-0.515	-0.362	-0.227	-0.090
7.6	-2.291	-1.604	-1.202	-0.917	-0.696	-0.515	-0.362	-0.227	-0.091
7.7	-2.291	-1.604	-1.202	-0.917	-0.696	-0.515	-0.362	-0.227	-0.092
7.8	-2.292	-1.604	-1.202	-0.917	-0.696	-0.515	-0.362	-0.227	-0.092
7.9	-2.292	-1.604	-1.202	-0.917	-0.696	-0.515	-0.362	-0.227	-0.093
8.0	-2.292	-1.605	-1.202	-0.917	-0.696	-0.515	-0.362	-0.227	-0.093
8.1	-2.292	-1.605	-1.203	-0.917	-0.696	-0.514	-0.361	-0.227	-0.094
8.2	-2.293	-1.605	-1.203	-0.917	-0.696	-0.514	-0.361	-0.227	-0.094
8.3	-2.293	-1.605	-1.203	-0.917	-0.695	-0.514	-0.361	-0.227	-0.095
8.4	-2.293	-1.605	-1.203	-0.917	-0.695	-0.514	-0.361	-0.227	-0.095
8.5	-2.293	-1.605	-1.203	-0.917	-0.695	-0.514	-0.361	-0.227	-0.096
8.6	-2.293	-1.605	-1.203	-0.917	-0.695	-0.514	-0.361	-0.227	-0.096
8.7	-2.294	-1.605	-1.203	-0.917	-0.695	-0.514	-0.361	-0.227	-0.096
8.8	-2.294	-1.606	-1.203	-0.917	-0.695	-0.514	-0.361	-0.227	-0.097
8.9	-2.294	-1.606	-1.203	-0.917	-0.695	-0.514	-0.361	-0.227	-0.097
9.0	-2.294	-1.606	-1.203	-0.917	-0.695	-0.514	-0.361	-0.227	-0.098
9.1	-2.294	-1.606	-1.203	-0.917	-0.695	-0.514	-0.360	-0.227	-0.098
9.2	-2.295	-1.606	-1.203	-0.917	-0.695	-0.514	-0.360	-0.227	-0.098
9.3	-2.295	-1.606	-1.203	-0.917	-0.695	-0.514	-0.360	-0.227	-0.098
9.4	-2.295	-1.606	-1.203	-0.917	-0.695	-0.514	-0.360	-0.227	-0.099
9.5	-2.295	-1.606	-1.203	-0.917	-0.695	-0.514	-0.360	-0.227	-0.099
9.6	-2.296	-1.606	-1.203	-0.917	-0.695	-0.513	-0.360	-0.227	-0.099
9.7	-2.296	-1.606	-1.203	-0.917	-0.695	-0.513	-0.360	-0.227	-0.100
9.8	-2.296	-1.606	-1.203	-0.917	-0.695	-0.513	-0.360	-0.227	-0.100
9.9	-2.296	-1.606	-1.203	-0.917	-0.695	-0.513	-0.360	-0.227	-0.100
10	-2.296	-1.606	-1.203	-0.917	-0.695	-0.513	-0.360	-0.227	-0.100
11	-2.296	-1.606	-1.203	-0.917	-0.695	-0.512	-0.360	-0.227	-0.100
12	-2.296	-1.606	-1.203	-0.917	-0.694	-0.512	-0.360	-0.226	-0.100
13	-2.296	-1.606	-1.203	-0.917	-0.694	-0.511	-0.360	-0.225	-0.100
14	-2.296	-1.606	-1.203	-0.917	-0.694	-0.511	-0.358	-0.225	-0.100
15	-2.299	-1.608	-1.203	-0.916	-0.693	-0.511	-0.358	-0.224	-0.100
20	-2.300	-1.608	-1.203	-0.916	-0.693	-0.511	-0.357	-0.224	-0.100
50	-2.302	-1.609	-1.203	-0.916	-0.693	-0.510	-0.356	-0.224	-0.105
100	-2.302	-1.609	-1.203	-0.916	-0.693	-0.510	-0.356	-0.224	-0.105
1000	-2.302	-1.609	-1.203	-0.916	-0.693	-0.510	-0.356	-0.224	-0.105
10000	-2.302	-1.609	-1.203	-0.916	-0.693	-0.510	-0.356	-0.224	-0.105

TABLE 7

Size of the test of the hypothesis that the data are normal / exponential

$$\boldsymbol{\lambda = \theta/\sigma_v = 1}$$

Nominal size = 0.05

k	*n*	Wrong Skew (%)	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
3	50	47.49	0.075	0.050	0.043	0.039
	100	40.55	0.065	0.045	0.044	0.043
	250	38.17	0.054	0.051	0.048	0.045
5	50	47.49	0.080	0.039	0.041	*
	100	40.55	0.067	0.049	0.044	*
	250	38.17	0.055	0.046	0.048	*
10	50	47.49	0.179	0.038	0.040	*
	100	40.55	0.116	0.042	0.046	*
	250	38.17	0.069	0.052	0.051	*

* The number of cells (k) is not relevant for the Kolmogorov-Smirnov test.

TABLE 8

Power of the test of the hypothesis that the data are normal / exponential

Alternative: the data are normal / half normal (σ_u^2)

Nominal size = 0.05

Number of cells: k = 3

n	σ_u	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
50	0.1	0.050	0.033	0.043	0.029
	0.5	0.048	0.030	0.036	0.033
	1	0.059	0.040	0.041	0.045
	2	0.071	0.051	0.054	0.039
	5	0.105	0.082	0.069	0.058
	10	0.138	0.101	0.094	0.093
100	0.1	0.088	0.042	0.039	0.036
	0.5	0.074	0.044	0.050	0.041
	1	0.049	0.049	0.061	0.040
	2	0.064	0.052	0.059	0.047
	5	0.145	0.131	0.122	0.162
	10	0.263	0.201	0.220	0.248
250	0.1	0.066	0.055	0.054	0.042
	0.5	0.060	0.061	0.049	0.043
	1	0.063	0.068	0.052	0.043
	2	0.071	0.068	0.050	0.054
	5	0.289	0.271	0.263	0.422
	10	0.593	0.553	0.530	0.682

TABLE 9A

Power of the test of the hypothesis that the data are normal / exponential

Alternative: the data are normal / gamma (*u* is *c* times gamma(*m*))

Nominal size = 0.05

***m* = 0.5**

n	*c*	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
50	0.1	0.043	0.051	0.052	0.043
	0.5	0.060	0.044	0.050	0.045
	1	0.087	0.060	0.055	0.054
	2	0.067	0.062	0.049	0.051
	5	0.119	0.087	0.077	0.127
	10	0.218	0.162	0.155	0.282
100	0.1	0.053	0.045	0.045	0.039
	0.5	0.063	0.051	0.049	0.045
	1	0.065	0.049	0.049	0.064
	2	0.087	0.070	0.062	0.086
	5	0.192	0.180	0.181	0.282
	10	0.352	0.334	0.321	0.507
250	0.1	0.077	0.070	0.067	0.052
	0.5	0.081	0.080	0.077	0.053
	1	0.072	0.079	0.070	0.066
	2	0.090	0.075	0.067	0.069
	5	0.416	0.395	0.384	0.467
	10	0.789	0.765	0.752	0.880

TABLE 9B

Power of the test of the hypothesis that the data are normal / exponential

Alternative: the data are normal / gamma (*u* is *c* times gamma(*m*))

Nominal size = 0.05

***m* = 2**

n	*c*	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
50	0.1	0.036	0.030	0.033	0.025
	0.5	0.056	0.041	0.039	0.034
	1	0.056	0.036	0.040	0.031
	2	0.073	0.049	0.057	0.045
	5	0.101	0.080	0.081	0.065
	10	0.129	0.112	0.106	0.124
100	0.1	0.055	0.048	0.051	0.037
	0.5	0.085	0.065	0.061	0.045
	1	0.062	0.068	0.059	0.045
	2	0.076	0.070	0.070	0.065
	5	0.115	0.093	0.088	0.109
	10	0.150	0.135	0.133	0.176
250	0.1	0.065	0.060	0.056	0.056
	0.5	0.079	0.069	0.072	0.061
	1	0.095	0.087	0.079	0.071
	2	0.098	0.086	0.088	0.079
	5	0.199	0.181	0.171	0.278
	10	0.292	0.275	0.281	0.372

REFERENCES

Abadir, K.M. and J.R. Magnus (2005), *Matrix Algebra* (Econometric Exercises, Volume 1), Cambridge, Cambridge University Press.

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

Bai, J. (2003), "Testing Parametric Conditional Distributions of Dynamic Models," *The Review of Economics and Statistics* 85, 531-549.

Bera, A.K. and N.C. Mallick (2002), "Information Matrix Tests for the Composed Error Frontier Model," Chapter 32 in *Advances on Methodological and Applied Aspects of Probability and Statistics*, N. Balakrishnan, editor, Gordon and Breach Science Publishers.

Chen, Y.-T. and H.-J. Wang (2009), "Centered-Residuals-Based Moment Estimator and Test for Stochastic Frontier Models," unpublished manuscript, Academia Sinica.

Coelli, T. (1995), "Estimators and Hypothesis Tests for a Stochastic Frontier Function," *Journal of Productivity Analysis* 6, 247-265.

Coelli, T., D.S. Prasada Rao, C.J. O'Donnell and G.E. Battese (2005), *An Introduction to Efficiency and Productivity Analysis*, 2nd edition, New York, Springer.

Giné, E. and J. Zinn (1990), "Bootstrapping General Empirical Measures," *Annals of Probability* 18, 851-869.

Greene, W.H. (1980a), "Maximum Likelihood Estimation of Econometric Frontier Functions," *Journal of Econometrics* 13, 27-56.

Greene, W.H. (1980b), "On the Estimation of a Flexible Frontier Production Model," *Journal of Econometrics* 13, 101-115.

Greene, W.H. (1990), "A Gamma-Distributed Stochastic Frontier Model," *Journal of Econometrics* 46, 141-164.

Greene, W.H. (2008), *Econometric Analysis,* 6th edition, Upper Saddle River, N.J., Pearson Prentice Hall.

Hansen, L.P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica* 50, 1029-1054.

Heckman, J. (1984), "The χ^2 Goodness of Fit for Models Estimated from Microdata," *Econometrica* 52, 1543-1548.

Johnson, N.L. and S. Kotz (1970), *Continuous Univariate Distributions* – 1, Boston, Houghton Mifflin.

Jondrow, J., C.A.K. Lovell, I.S. Materov, and P. Schmidt (1982), "On the Estimation of Technical Efficiency in the Stochastic Frontier Production Function Model," *Journal of Econometrics* 19, 233-238.

Khmalzade, E.V. (1981), "Martingale Approach to the Theory of Goodness of Fit Test," *Theory of Probability and Its Applications*, 26, 240-257.

Khmalzade, E.V. (1988), "An Innovation Approach in Goodness of Fit Tests in R^m," *Annals of Statistics* 16, 1503-1516.

Khmalzade, E.V. (1993), "Goodness of Fit Problem and Scanning Innovation Martingales," *Annals of Statistics* 21, 798-829.

Kopp, R.J. and J. Mullahy (1990), "Moment-Based Estimation and Testing of Stochastic Frontier Models," *Journal of Econometrics* 46, 165-183.

Lee, L.-F. (1983), "A Test for Distributional Assumptions for the Stochastic Frontier Functions," *Journal of Econometrics* 22, 245-267.

Meeusen, W., and J. van den Broeck (1977), "Efficient Estimation from Cobb-Douglas Production Functions with Composed Error," *International Economic Review* 18, 435-444.

Newey, W.K. (1985), "Maximum Likelihood Specification Testing and Conditional Moment Tests," *Econometrica* 53, 1047-1070.

Pitt, M.M., and L.F. Lee (1981), "The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry," *Journal of Development Economics* 9, 43-64.

Ruppert, D. and R.J. Carroll (1980), "Trimmed Least Squares Estimation in the Linear Model," *Journal of the American Statistical Association* 75, 828-838.

Schmidt, P. and T.-F. Lin (1984), "Simple Tests for Alternative Specifications in Stochastic Frontier Models," *Journal of Econometrics* 24, 349-361.

Simar, L. and P.W. Wilson (2010), "Inferences from Cross-Sectional Stochastic Frontier Models," *Econometric Reviews* 29, 62-98.

Stevenson, R.E. (1980), "Likelihood Functions for Generalized Stochastic Frontier Estimation," *Journal of Econometrics* 13, 57-66.

Stute, W., W. Gonzáles Manteiga and M. Presedo Quindimil (1993), "Bootstrap Based Goodness of Fit Tests," *Metrika* 40, 243-256.

Tauchen, G. (1985), "Diagnostic Testing and Evaluation of Maximum Likelihood Models," *Journal of Econometrics* 30, 415-444.

Tallis, G.M. (1983), "Goodness of Fit," pp. 451-461 in S. Kotz and N.L. Johnson, *Encyclopedia of Statistical Sciences,* Vol. 3, New York, Wiley.

Waldman, D. (1982), "A Stationary Point for the Stochastic Frontier Likelihood," *Journal of Econometrics* 18, 275-279.

Wang, W.S. and P. Schmidt (2009), "On the Distribution of Estimated Technical Efficiency in Stochastic Frontier Models," *Journal of Econometrics* 148, 36-45.

White, H. (1982), "Maximum Likelihood Estimation of Misspecified Models," *Econometrica* 50, 1-16.

Zellner, A. and N. Revankar (1970), "Generalized Production Functions," *Review of Economic Studies* 37, 241-250.