PROPERTIES OF TECHNICAL EFFICIENCY ESTIMATORS IN
THE STOCHASTIC FRONTIER MODEL

Donald M. WALDMAN*
University of North Carolina, Chapel Hill, NC 27514, USA
Received June 1983, final version received December 1983

This paper considers the disturbance specification \( \varepsilon = \nu - u \) of the stochastic frontier model. For \( \nu \) distributed zero-mean normal and \( u \) half normal or exponential, we evaluate the population correlation coefficients between \( u \) and three estimators of \( u \), \( \text{E}(u|\varepsilon) \) and two linear estimators, for various values of the signal-to-noise ratio.

1. Introduction

The estimation of frontier cost and production functions has been given impetus by the introduction of the stochastic frontier.\(^1\) This model is distinguished by a disturbance term composed of two parts - a one-sided component, non-positive for production functions and non-negative for cost functions, representing the degree of inefficiency, and a symmetric component representing the usual statistical noise (reporting and measurement error, for example) that characterizes any relationship. This formulation has proven to alleviate some of the statistical and practical shortcomings of earlier attempts at measuring frontiers, such as linear programming methods.

A serious drawback of this procedure is the inability to measure individual effects. This was the original motivation for the introduction of production isoquants in the pioneering work of Farrell (1957). Instead, all that is available is an estimate of the overall or industry level of inefficiency. In an attempt to rectify this shortcoming, Jondrow et al. (1982) proposed estimating the firm-level inefficiency with an estimate of the conditional expectation of the one-sided component conditional on the total disturbance. This is an interesting suggestion for an estimator, but the account of its statistical properties is limited to the observation that the sampling error disappears asymptotically. Since the total disturbance contains only imperfect information regarding the one-sided disturbance, the proposed estimator has a certain 'intrinsic' variability [Jondrow et al. (1982, p. 235)].

*I am grateful to Randy Nelson, Peter Schmidt, V. Kerry Smith, and an anonymous referee for helpful suggestions.

\(^1\)See Aigner et al. (1977) and Meeusen and van den Broeck (1977).

0304-4076/84/$3.00 ©1984, Elsevier Science Publishers B.V. (North-Holland)
In this paper we discuss the statistical justification for the proposed estimator, and deal explicitly with the extent of the intrinsic variability. The conditional expectation function is unbiased and, in general, non-linear, so that a natural comparison of the estimator would be with linear estimation or prediction. We consider both a commonly applied linear unbiased estimator and best linear prediction. The question arises as to which estimator is preferred. The answer will depend upon the use to which the final results are put. Estimates of firm-level inefficiency may be used to rank a sample of firms. For this purpose the conditional expectation estimator and the two linear estimators will produce the identical ordering. Alternatively, the relative degree of inefficiency among firms may be important. Below we prove a general inequality between correlation coefficients which suggests that the conditional expectation estimator is always preferred to either linear estimator on this basis. The two linear estimators correlate equally well in the population with the inefficiency. Finally, an absolute measure of inefficiency for each firm may be desired. For this purpose, conditional expectation estimation is preferred to best linear prediction, which in turn is preferred to unbiased linear estimation. The criterion here is mean-squared error, which is equivalent to the variance of the prediction error [see Theil (1971, p. 123)].

The inequalities mentioned above will be strict except for trivial cases (correlations of zero or one). This is not the full story, however. The important question remains as to the extent of the gain from using the preferred estimator. This is examined below by calculating the population correlation coefficients between the estimators and the inefficiency, and the variance of the prediction error for each estimator. Consideration of the results is simplified by the fact that these correlations depend only upon the signal-to-noise ratio in the disturbance. The complexity of the production process is irrelevant, and since we deal with the population, there is no sample size factor. Results are presented for the normal plus half-normal and and normal plus exponential specifications, those considered by Jondrow et al. (1982), and most commonly assumed in the applied literature.

Estimating technical inefficiency is a difficult venture, but not one that economists have shied away from (witness the recent examples in diverse areas that have appeared in the literature). Future developments are presaged by the renewed interest in index numbers, incipient work in non-parametric methods, and the use of systems of equations (cost shares or input demands) as well as richer data sets (multiple observations on the same firms). Some or all of these developments may prove superior to the existing estimators compared here. But the fact remains that the predominant method of estimating inefficiency is to estimate an average or full frontier (by least squares) or a stochastic frontier on a single cross-section of firms. The results contained in this paper will be of use to those involved in such a study, as well as to those evaluating the attempts to extend the existing methodology.
In the next section the disturbance specification of the stochastic frontier model is reviewed and the estimators introduced. Then a lemma concerning correlation coefficients is proven. Calculation of the correlations is rather complicated and their derivations are relegated to appendix A. There the relationship between the correlation and the signal-to-noise ratio is shown; additional results appear in appendix B. These results involve expectations of the Mill’s ratio and functions of Mill’s ratio and its argument, and may well be of general interest.

Table 1 presents population correlation coefficients for the correlation between each of the three proposed estimators and the inefficiency, for various values of the signal-to-noise ratio. In section 5 a second lemma is proven, which substantiates the claim made above about the relative merits of the three estimators with respect to a measure of absolute inefficiency. Table 2 presents the root mean-squared errors (or the standard deviation of the prediction error) for the three estimators. Section 6 concludes.

2. Specification and alternative estimators

The disturbance in the stochastic frontier model has the form

$$\varepsilon = v - u,$$

where \( v \sim N(0, \sigma_v^2) \) is the symmetric component. Two stochastic assumptions are commonly made for the one-sided component \( u \): the half-normal distribution and the exponential distribution. Jondrow et al. (1982) propose using an estimate of the conditional expectation of \( u \) given \( \varepsilon \), \( E(u|\varepsilon) \), as an estimator of \( u \). This is a non-linear function of \( \varepsilon \), and is an unbiased estimator of \( u \) in the Theil (1971, p. 183) sense of zero expected prediction error since \( E[E(u|\varepsilon)] = E(u) \). To obtain estimates of \( E(u|\varepsilon) \), residuals from estimating the stochastic frontier model are used as inputs into the non-linear function.

Two linear estimators are considered. The first is simply \(-\varepsilon\), which is justified by the fact that \( E(-\varepsilon) = E(-v + u) = E(u) \). A more important reason for considering this estimator is that often the random disturbance \( (v) \) is ignored and a ‘full’ frontier is fit to the data. In the production function case this means that no observation may lie above the frontier. One method of obtaining firm-specific measures of inefficiency is to estimate by least squares and subtract the largest (positive) residual from each residual in the sample – see Greene (1980). The properties of this estimator will be identical

---

2 These residuals could be calculated simply by the ‘COLS’ method, which utilizes least squares slope coefficients estimators and adjusts the intercept by a constant (the estimated expectation of \( u \)). Olsen et al. (1980) conclude that this estimation method is actually preferred to maximum likelihood estimation for sample sizes below 400. Since residuals calculated in this way represent a simple translation of least squares residuals, the comparison with the linear estimators is exact.
D. M. Waldman, Properties of technical efficiency estimators

to the properties of the estimator $-\varepsilon$. We will refer to $-\varepsilon$ as ‘linear unbiased estimation’.

The second linear estimator to be considered is $\alpha + \beta \varepsilon$, where

$$\beta = -\frac{V(u)}{V(u) + V(\varepsilon)},$$

and

$$\alpha = E(u) - \beta E(\varepsilon) = E(u)(1 + \beta).$$

This estimator is also unbiased, since

$$E[E(u)(1 + \beta) + \beta \varepsilon] = E(u)(1 + \beta) - \beta E(u) = E(u).$$

Since it is formulated under the principles of best linear estimation, it will be preferred to $-\varepsilon$ by the mean-squared error criterion. We will refer to $\alpha + \beta \varepsilon$ as ‘best linear prediction’. Stochastic frontier residuals may be used to estimate $\varepsilon$, while the additional parameters in $\alpha$ and $\beta$ [and in $E(u|\varepsilon)$] are the natural by-produces of stochastic frontier estimation.

3. An inequality among correlations

That the conditional expectation suggestion has merit is shown by:

Lemma 1. For any two random variables $X$ and $Y$,

$$|\rho_{Y, E(Y|X)}| \geq |\rho_{Y, X}|. \quad (2)$$

Proof. Without loss of generality, assume $E(X) = E(Y) = 0$. Let $Z = E(Y|X)$. Then $|\rho_{Y, Z}| = |E(YZ)|/\sigma_Y \sigma_Z$. But $E(YZ) = E_x[E(YZ|X)] = E_x[E(ZE(Y|X))] = E_x(Z^2)$, where the first equality is based on the ‘law of iterated expectations’, the second on the fact that $Z$ is constant given $X$, and the third on the definition of $Z$. Here the subscript indicates that expectation is taken over $X$. Since $E(Z) = E_x[E(Y|X)] = E(Y) = 0$, $E_x(Z^2) = \sigma_Z^2$. Hence

$$|\rho_{Y, Z}| = \sigma_Z^2/(\sigma_Y \sigma_Z) \leq \sigma_Z/\sigma_Y. \quad (3)$$

Now

$$|\rho_{Y, X}| = |E(Y, X)|/\sigma_Y \sigma_X = |E(X, Z)|/\sigma_X \sigma_X \leq \sigma_X \sigma_Z/\sigma_Y \sigma_Y = |\rho_{Y, Z}|,$$

where the second equality, again based on the law of iterated expectations, is followed by Schwartz’s inequality and substitution from (3). Q.E.D.

3A referee has pointed out that this ‘best linear’ estimator may have the wrong sign. A possible improvement, then, would be to use $\alpha + \beta \varepsilon$ if $\alpha + \beta \varepsilon > 0$ and 0 otherwise. But this would be a non-linear estimator inferior in theory to the conditional expectation function, the non-linear estimator considered here.
In the stochastic frontier model, this means that $E(u|\varepsilon)$ will always correlate with $u$ at least as well as $\varepsilon$. In fact, it is easy to show that the inequality is strict except for $\rho_{u,\varepsilon} = \rho_{u,E(u|\varepsilon)} = 0$ and $\rho_{u,\varepsilon} = \rho_{u,E(u|\varepsilon)} = 1$. Since $\alpha + \beta \varepsilon$ is simply a scaling and translation of $\varepsilon$, $\rho_{u,\alpha + \beta \varepsilon} = \rho_{u,\varepsilon}$. In the discussion below, therefore, we do not distinguish the two estimators.

We now turn to the question of the extent of the gain from using $E(u|\varepsilon)$ to estimate $u$ compared to the linear estimators.

4. Evaluation of the estimators

Let $\sigma_u^2$ be the single parameter describing the distribution of $u$. When $u \sim N(0, \sigma_u^2)$, $V(u) = ((\pi - 2)/\pi)\sigma_u^2$. Since the quantity $((\pi - 2)/\pi)$ appears throughout, define $\pi_* = (\pi - 2)/\pi$. When $u$ follows the exponential distribution, $V(u) = \sigma_u^2$. Let $\lambda^2 = \sigma_u^2/\sigma^2$ and $\lambda = \sigma_u/\sigma$. In appendix B it is shown that $\rho_{u,\varepsilon}$ and $\rho_{u,E(u|\varepsilon)}$ for both distributions of $u$ are functions only of $\lambda$. The calculating formulas for these quantities are derived in appendix A. Table 1 reports these calculations for selected values of $\lambda$.

Although Lemma 1 establishes the superiority of the conditional expectation over either linear estimator as to the amount of the information about the inefficiency, as the entries in table 1 indicate, the gain is not great. For the normal estimator, the maximum gain of 2.2% occurs where $\sqrt{V(u)/V(v)}$

<table>
<thead>
<tr>
<th>$\sqrt{V(u)/V(v)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10 0.20 0.30 0.40 0.50 0.60 0.70 0.80 0.90</td>
</tr>
<tr>
<td>$\rho_{u,E(u</td>
</tr>
<tr>
<td>$u \sim [N(0, \sigma_u^2)]$</td>
</tr>
<tr>
<td>0.100 0.198 0.292 0.381 0.461 0.532 0.594 0.646 0.691</td>
</tr>
<tr>
<td>$\rho_{u,\varepsilon}$</td>
</tr>
<tr>
<td>both models</td>
</tr>
<tr>
<td>0.100 0.196 0.287 0.371 0.447 0.514 0.573 0.625 0.669</td>
</tr>
<tr>
<td>$\rho_{u,E(u</td>
</tr>
<tr>
<td>$u \sim \exp(\sigma_u)$</td>
</tr>
<tr>
<td>$^a$</td>
</tr>
<tr>
<td>0.308 0.403 0.501 0.578 0.641 0.692 0.733</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\sqrt{V(u)/V(v)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00 1.25 1.50 1.75 2.00 3.00 5.00 10.00</td>
</tr>
<tr>
<td>$\rho_{u,E(u</td>
</tr>
<tr>
<td>$u \sim [N(0, \sigma_u^2)]$</td>
</tr>
<tr>
<td>0.729 0.800 0.848 0.881 0.904 0.953 0.982 0.995</td>
</tr>
<tr>
<td>$\rho_{u,\varepsilon}$</td>
</tr>
<tr>
<td>both models</td>
</tr>
<tr>
<td>0.707 0.781 0.832 0.868 0.894 0.949 0.981 0.995</td>
</tr>
<tr>
<td>$\rho_{u,E(u</td>
</tr>
<tr>
<td>$u \sim \exp(\sigma_u)$</td>
</tr>
<tr>
<td>0.767 0.829 0.869 0.897 0.917 0.958 0.983 0.995</td>
</tr>
</tbody>
</table>

$^a$ Omitted because the calculation was deemed inaccurate. See footnote 4.
equals 0.90 and 1.00. In the exponential case, the maximum gain is larger, 6.8%, occurring at \( \sqrt{V(u)/V(v)} = 0.70 \). In both cases the advantage diminishes in either direction. It should be noted that comparisons should be made only between \( \rho_{U,E(u|e)} \) for the normal model and \( \rho_{u,e} \), and between \( \rho_{U,E(u|e)} \) for the exponential model and \( \rho_{u,e} \), and not between the correlation coefficients for the two stochastic specifications. The reason for the non-comparability, of course, is precisely that the underlying models are different – the two correlations are not competing estimators for the same model. Ideally, a consideration of the economic model should dictate the choice of specification. Both the half-normal and the exponential distributions are plausible choices for \( u \), as each describes a non-negative random variable with monotonically decreasing density. In practice, there may be no compelling reasons for choosing between the two, or other candidate distributions. In the absence of any other information, the exponential model might be chosen on the basis of its uniformly higher correlations, but in any event the gain here is not great, taking a maximum of 4.7% at \( \sqrt{V(u)/V(v)} = 0.70 \).

5. The variance of the prediction error

Lemma 1 of section 3 may be easily translated into an analogous inequality regarding expected squared loss. For this criterion we need an additional result regarding the linear unbiased estimator. Without loss of generality, we consider the zero-mean case:

**Lemma 2.** For any two random variables \( X \) and \( Y \) such that \( E(X) = E(Y) = 0 \), with \( \beta = \sigma_{xy}/\sigma_x^2 \), we have

\[
E(Y - X)^2 \geq E(Y - \beta X)^2 \geq E(Y - E(Y|X))^2. \tag{4}
\]

**Proof.** For the first inequality, write \( Y - X = Y - \beta X + X(\beta - 1) \). Then the inequality is easily verified by squaring both sides and noting that \( E[X(Y - \beta X)] = 0 \). For the frontier stochastic specification of eq. (1), by analogy we have \( \sigma_{xy} = \sigma_y^2 \). Then

\[
E(Y - X)^2 = \sigma_x^2 - \sigma_y^2 = \sigma_y^2(\sigma_x^2/\sigma_y^2 - 1).
\]

Also under this specification

\[
\rho_{x,y}^2 = \sigma_{xy}^2/\sigma_x^2 \sigma_y^2 = \sigma_y^2/\sigma_x^2,
\]

so that

\[
E(Y - X)^2 = \sigma_y^2(1 - \rho_{x,y}^2) \rho_{x,y}^2. \tag{5}
\]

To prove the second inequality in the general case and for comparison with (5)
in the particular case at hand we proceed analogously:

\[ E(\mathbf{Y} - \mathbf{B}\mathbf{X})^2 = \sigma_y^2 + \beta^2 \sigma_x^2 - 2\beta \sigma_{xy} = \sigma_y^2 - \beta^2 \sigma_x^2, \]

by the definition of \( \beta \). But

\[
1 - \rho^2_{x,y} = 1 - \sigma^2_{xy}/\sigma_x^2 = 1 - \beta \sigma_{xy}/\sigma_y^2 = (1/\sigma_y^2) \left( \sigma_y^2 - \beta \sigma_{xy} \right)
\]

\[ = \left(1/\sigma_y^2 \right) \left( \sigma_y^2 - \beta^2 \sigma_x^2 \right), \]

so that

\[ E(\mathbf{Y} - \mathbf{B}\mathbf{X})^2 = \sigma_y^2 \left(1 - \rho^2_{x,y} \right). \tag{6} \]

For the third quantity in (4), again let \( \mathbf{Z} = E(\mathbf{Y}|\mathbf{X}) \). Then

\[ E(\mathbf{Y} - \mathbf{Z})^2 = \sigma_y^2 + E(\mathbf{Z}^2) - 2E(\mathbf{YZ}) = \sigma_y^2 - \sigma_z^2, \]

(see the proof of Lemma 1). But

\[ 1 - \rho^2_{y,z} = 1 - \sigma^2_y/\sigma_z^2 = \left(1/\sigma_y^2 \right) \left( \sigma_y^2 - \sigma_z^2 \right), \]

Table 2

| \( \sqrt{\text{V}(\mu)/\text{V}(\nu)} \) |
|---|---|---|---|---|---|---|---|---|
| 0.10 | 0.20 | 0.30 | 0.40 | 0.50 | 0.60 | 0.70 | 0.80 |
| \( E(\mu|\epsilon) \) | | | | | | | |
| \( u \sim [N(0, \sigma_u^2)] \) | 0.995 | 0.980 | 0.956 | 0.925 | 0.887 | 0.847 | 0.804 | 0.763 |
| \( \alpha + \beta \epsilon \) | | | | | | | |
| both models | 0.995 | 0.981 | 0.958 | 0.929 | 0.895 | 0.858 | 0.820 | 0.781 |
| \( \epsilon \) | | | | | | | |
| both models | 9.95 | 5.00 | 3.34 | 2.50 | 2.00 | 1.67 | 1.43 | 1.25 |
| \( E(\mu|\epsilon) \) | | | | | | | |
| \( u \sim \exp(\sigma_u) \) | | | | | | | |

| \( \sqrt{\text{V}(\mu)/\text{V}(\nu)} \) |
|---|---|---|---|---|---|---|---|---|
| 0.90 | 1.00 | 1.25 | 1.50 | 1.75 | 2.00 | 3.00 | 5.00 | 10.00 |
| \( E(\mu|\epsilon) \) | | | | | | | |
| \( u \sim [N(0, \sigma_u^2)] \) | 0.723 | 0.685 | 0.600 | 0.530 | 0.473 | 0.428 | 0.303 | 0.189 | 0.100 |
| \( \alpha + \beta \epsilon \) | | | | | | | |
| both models | 0.743 | 0.707 | 0.625 | 0.555 | 0.497 | 0.448 | 0.315 | 0.194 | 0.100 |
| \( \epsilon \) | | | | | | | |
| both models | 1.11 | 1.00 | 0.801 | 0.667 | 0.572 | 0.501 | 0.332 | 0.198 | 0.100 |
| \( E(\mu|\epsilon) \) | | | | | | | |
| \( u \sim \exp(\sigma_u) \) | 0.680 | 0.642 | 0.559 | 0.495 | 0.442 | 0.399 | 0.287 | 0.184 | 0.100 |

\( ^a \)Omitted because the calculation was deemed inaccurate. See footnote 4.
so that
\[ E(Y - Z)^2 = \sigma^2_y (1 - \rho^2_{y,z}). \]  

(7)

Now comparing (6) and (7) and using the results of Lemma 1 proves the second inequality in (4). Eqs. (5)-(7) and table 1 may be used to construct table 2, which compares the ratios of root-mean-squared errors to standard deviations \[ \sqrt{E(Y - f(X))^2}/\sigma_y \] for various \( f(X) \).

As is clear from eq. (5) and table 2, the mean-squared error of \(-\varepsilon\) deteriorates badly as the signal-to-noise ratio approaches zero, as it should. Recall that this is the estimator that ignores the existence of noise.

6. Concluding remarks

We have examined three alternative estimators of micro-level inefficiency in the estimation of the stochastic frontier model: the conditional expectation function, a linear unbiased estimator that ignores the stochastic nature of the frontier, and best linear prediction. With respect to ranking a cross-section of firms as to their degree of inefficiency, the three estimators are identical. The conditional expectation is preferred to either linear estimator (both of which produce the same result) with respect to measuring the relative inefficiency of the firms. This is because the conditional expectation estimator takes advantage of the form of the distribution function. If the exponential model is assumed, the gain is somewhat greater than when the normal model is assumed, but in either event the advantage must be deemed marginal. This should come as no surprise considering the difficult task of decomposing an unobservable variable into the sum of two unobservable variables.

With respect to the variance of the prediction error or the mean-squared error of the estimators, the conditional expectation is preferred to the best linear predictor, but again the difference is small. For this criterion, the best linear predictor dominates the linear unbiased estimator, which for small values of the signal-to-noise ratio has a large mean-squared error.

Appendix A: Derivation and calculation of the correlations

A few preliminary results will ease the notation. Let \( \phi(a) = f(a)/F(a) \), the ratio of the standard normal density function to the standard normal cumulative distribution function. From Johnson and Kotz (1970, pp. 81–82), we have the following formulas: if \( X \sim N(\mu, \sigma^2) \), \( b = \mu/\sigma \), then

\[ E(X) = \mu + \sigma (f(-b)/1 - F(-b)) = \sigma (\phi(b) + b), \]  

(A.1)

and

\[ V(X) = \sigma^2 [1 - b\phi(b) - \phi^2(b)], \]  

(A.2)
A basic variance decomposition is, for any two random variables \( u \) and \( \varepsilon \),
\[
V(u) = V_E(u|\varepsilon) + E_v V(u|\varepsilon),
\]
(A.3)
where the subscript \( \varepsilon \) indicates expectation is taken over values of \( \varepsilon \). From eq. (3), the squared correlation coefficient of \( u \) with \( E(u|\varepsilon) \) is
\[
\rho_{u,E(u|\varepsilon)}^2 = \frac{\sigma_{E(u|\varepsilon)}^2}{\sigma_u^2}.
\]
(A.4)
Substituting (A.3) into (A.4) yields
\[
\rho_{u,E(u|\varepsilon)}^2 = \frac{V(u) - EV(u|\varepsilon)}{V(u)} = 1 - EV(u|\varepsilon)/V(u).
\]
(A.5)

Now there are two cases. First, for \( u \sim N(0, \sigma_u^2) \), Jondrow et al. (1982) give \( u|\varepsilon \sim N(\mu_*, \sigma_\varepsilon^2) \), where \( \mu_* = -\sigma_u \varepsilon / \sigma^2 \), \( \sigma_*^2 = \sigma_u^2 \sigma_\varepsilon^2 / \sigma^2 \), and \( \sigma^2 = \sigma_u^2 + \sigma_\varepsilon^2 \). Therefore, as in (A.1) and (A.2), let \( b = \mu_*/\sigma_* = -\lambda \varepsilon / \sigma \), with \( \lambda = \sigma_u / \sigma_\varepsilon \), then (A.2) specializes to \( V(u|\varepsilon) = \sigma_*^2 (1 - b \phi(b) - \phi^2(b)) \), and with \( b = 0 \), (A.2) specializes to \( V(u) = \sigma_u^2 (1 - 2/\pi) = \pi \sigma_u^2 \), since \( f(0) = 1/\sqrt{2\pi} \) and \( F(0) = 0.5 \). Now, in appendix B below we show
\[
E[\phi^2(b)] = 0 \quad \text{where} \quad b = -\lambda \varepsilon / \sigma.
\]
(A.6)
Substituting (A.6) and the results above (A.6) into (A.5) yields
\[
\rho_{u,E(u|\varepsilon)}^2 = 1 - \frac{\sigma_*^2 [1 - E(\phi^2(b))] / \sigma_* \sigma_u^2}{\sigma_u^2}.
\]
(A.7)
Now \( \sigma_*^2 / \sigma_u^2 = (\sigma_u^2 \sigma_\varepsilon^2 / \sigma^2) / \sigma_u^2 = \sigma_\varepsilon^2 / (\sigma_u^2 + \sigma_\varepsilon^2) = 1/(1 + \lambda^2) \), so that (A.7) may be written
\[
\rho_{u,E(u|\varepsilon)}^2 = 1 - \frac{1 - E(\phi^2(b))}{\sigma_* (1 + \lambda^2)}.
\]
(A.8)
Eq. (A.8) was used to calculate \( \rho_{u,E(u|\varepsilon)} \) in table 1. There does not appear to be a closed form expression for \( E[\phi^2(b)] \), so that this term was evaluated by high precision numerical quadrature.\(^4\) In appendix B it is shown that this term is a function of \( \lambda \) only.

The second case is when \( u \) has the exponential distribution, with density \( g(u) = \exp(-u/\sigma_u) / \sigma_u \). Then Jondrow et al. (1982) give \( u|\varepsilon \sim N(\sigma \varepsilon, \sigma_u^2) \), where \( b = -\varepsilon / \sigma_\varepsilon - \lambda^{-1} \). Hence (A.2) becomes \( V(u|\varepsilon) = \sigma_\varepsilon^2 [1 - b \phi(b) - \phi^2(b)] \), and \( V(u) = \sigma_\varepsilon^2 \). In appendix B we show
\[
E[\phi^2(b)] = -\lambda^{-2}.
\]
(A.9)
\(^4\) We used 20-point Hermite polynomial quadrature. Fortran-coded subroutines are available from the author at cost.
Substituting (A.9) and the results above (A.9) into (A.5) yields

\[ \rho_{u, E(u|e)}^2 = 1 - \left( \frac{\sigma_u^2}{\sigma_v^2} \right) \left[ 1 + \lambda^{-2} - \text{E}\left( \varphi^2(b) \right) \right] \]

(A.10)

\[ = 1 - \lambda^{-2} \left[ 1 + \lambda^{-2} - \text{E}\left( \varphi^2(b) \right) \right]. \]

Eq. (A.10) was used to calculate \( \rho_{u, E(u|e)}^2 \) in table 1. Again there does not appear to be a closed form expression for \( \text{E}\left[ \varphi^2(b) \right] \) so that numerical methods were used, but as in the half-normal case it is easy to show that this term depends only upon \( \lambda \) (see appendix B).

For \( \rho_{u, \epsilon}^2 \) in the normal model multiply \( \epsilon = \nu - \mu \) by \( \mu - \text{E}(\mu) \) and take expectations giving \( \text{Cov}^2(u, \epsilon) = \text{V}(u)^2 = \pi^2 \sigma_u^4 \). Since \( \text{V}(\epsilon) = \pi^2 \sigma_u^2 + \sigma_\epsilon^2 \) and \( \text{V}(u) = \pi^2 \sigma_u^2 \), we have

\[ \rho_{u, \epsilon}^2 = \frac{\text{Cov}^2(u, \epsilon)}{\text{V}(u) \text{V}(\epsilon)} = \pi^2 \frac{\sigma_u^4}{(\pi^2 \sigma_u^2 + \sigma_\epsilon^2)} = \pi^2 \frac{\sigma_u^2}{(\pi^2 \sigma_u^2 + \sigma_\epsilon^2)} \]

(A.11)

Eq. (A.11) was used to calculate the entries in the row marked \( \rho_{u, \epsilon}^2 \) in table 1.

In order to make the results for the normal and exponential models comparable, the correlation coefficient should be calculated for constant \( \text{V}(u) / \text{V}(\nu) \). This quantity is equal to \( \pi^2 \lambda^2 \) in the normal model and \( \lambda^2 \) in the exponential model. In the exponential model, \( \text{Cov}^2(u, \epsilon) = \text{V}(u) = \sigma_u^2 \) and \( \text{V}(\epsilon) = \sigma_u^2 + \sigma_\epsilon^2 \).

Therefore

\[ \rho_{u, \epsilon}^2 = \frac{\sigma_u^4}{\sigma_u^2 (\sigma_u^2 + \sigma_\epsilon^2)} = \lambda^2 / (\lambda^2 + 1). \]

(A.12)

Considering (A.11) and the discussion above (A.12) we conclude that \( \rho_{u, \epsilon}^2 \) will be the same in each case, requiring only a single row in table 1.

Appendix B

**Proof of (A.6).** \( \text{E}[b \varphi(b)] = 0 \) where \( b = -\lambda \epsilon / \sigma \). The density of \( \epsilon \) is given by Aigner et al. (1977),

\[ g(\epsilon) = 2 F(b)f(\epsilon / \sigma) / \sigma = 2 F(b) f(b \lambda^{-1}) / \sigma. \]

(B.1)

Then

\[ \text{E}[b \varphi(b)] = (1 / \sigma) \int_{-\infty}^{\infty} b (f(b) / F(b)) \cdot 2 F(b) f(b \lambda^{-1}) \, d\epsilon \]

(B.2)

\[ = (2 / \sigma) \int_{-\infty}^{\infty} b f(b) f(b \lambda^{-1}) \, d\epsilon = 0, \]

since \( f(b) f(b \lambda^{-1}) \) is an even function of \( b \).
Proof of \( E[\phi^2(b)] = h(\lambda) \). From (B.1) above,
\[
E[\phi^2(b)] = \left(2/\sigma\right) \int_{-\infty}^{\infty} \left(f^2(b)/F^2(b)\right)f(\sigma/\epsilon)\,d\epsilon
\]
\[
= 2\int_{-\infty}^{\infty} \left(f^2(b)/F^2(b)\right)f(\sigma/\epsilon)(d\epsilon/\sigma).
\]

Let \( z = \epsilon/\sigma \). Then \( b = -\lambda \epsilon/\sigma = -\lambda z \) and \( dz = d\epsilon/\sigma \). Then
\[
E[\phi^2(b)] = 2\int_{-\infty}^{\infty} \left(f^2(-\lambda z)/F(-\lambda z)\right)f(z)\,dz = h(\lambda).
\]

Proof of (A.9). \( E[b\phi(b)] = -\lambda^{-2} \) where \( b = -\epsilon/\sigma_u - \lambda^{-1} \). The density of \( \epsilon \) in this case is given in Aigner et al. (1977),
\[
g(\epsilon) = F(b)\exp\left[\frac{1}{2}(\lambda^{-2} + 2\epsilon/\sigma_u)\right]/\sigma_u.
\]

Then
\[
E[b\phi(b)] = -E[\epsilon\phi(b)/\sigma_u + \lambda^{-1}\phi(b)] = -E[\lambda^{-1}\phi(b)],
\]

where the second equality sign in (B.6) is based on the same arguments used to prove (B.2). Using (B.5),
\[
E[\phi(b)] = \left(1/\sigma_u\right) \int_{-\infty}^{\infty} \left(f(b)/F(b)\right)F(b)\exp\left[\frac{1}{2}(\lambda^{-2} + 2\epsilon/\sigma_u)\right] \,d\epsilon
\]
\[
= \left(1/\sigma_u\right) \int_{-\infty}^{\infty} (1/\sqrt{2\pi})\exp\left[-\frac{1}{2}(\epsilon/\sigma_u - \lambda^{-1})^2\right]
\]
\[
\times \exp\left[\frac{1}{2}(\lambda^{-2} + 2\epsilon/\sigma_u)\right] \,d\epsilon
\]
\[
= \left(\sigma_u/\sigma_u\right) \int_{-\infty}^{\infty} (1/\sqrt{2\pi\sigma_v})\exp\left[-\frac{1}{2}(\epsilon^2/\sigma_v)\right] \,d\epsilon = \lambda^{-1}.
\]

Then, from (B.6), \( E[b\phi(b)] = -\lambda^{-2} \).

Proof of \( E[\phi^2(b)] = h(\lambda) \), exponential case. Using (B.5),
\[
E[\phi^2(b)] = \left(1/\sigma_u\right) \int_{-\infty}^{\infty} \left(f^2(b)/F(b)\right)\exp\left[\frac{1}{2}(\lambda^{-2} + 2\epsilon/\sigma_u)\right] \,d\epsilon.
\]
Let $z = \varepsilon / \sigma_u$. Then $b = -z - \lambda^{-1}$, $\varepsilon = \sigma_u z$, $\varepsilon / \sigma_u = \lambda^{-1} z$, and $dz = d\varepsilon / \sigma_u$. Then

$$E[\varphi^2(b)] = \lambda^{-1} \int_{-\infty}^{\infty} \left( f^2(-z - \lambda^{-1}) / F(-z - \lambda^{-1}) \right) \times \exp\left[\frac{1}{2}(\lambda^{-2} + 2\lambda^{-1} z)\right] dz = h(\lambda). \quad (B.9)$$

References


