Expected efficiency ranks from parametric stochastic frontier models

William C. Horrace · Seth Richards-Shubik · Ian A. Wright

Received: 13 April 2013 / Accepted: 12 January 2014 / Published online: 3 June 2014
© Springer-Verlag Berlin Heidelberg 2014

Abstract In the stochastic frontier model, we extend the multivariate probability statements of Horrace (J Econom, 126:335–354, 2005) to calculate the conditional probability that a firm is any particular efficiency rank in the sample. From this, we construct the conditional expected efficiency rank for each firm. Compared to the traditional ranked efficiency point estimates, firm-level conditional expected ranks are more informative about the degree of uncertainty of the ranking. The conditional expected ranks may be useful for empiricists. A Monte Carlo study and an empirical example are provided.

Keywords Efficiency estimation · Order statistics · Multivariate inference · Multiplicity

JEL Classification C12 · C16 · C44 · D24

1 Introduction

Given a sample of firm-level data, parametric stochastic frontier models specify production output (or cost) as the sum of a linear response function and an additively
composed error, consisting of a two-sided error, representing noise, and a one-sided error, representing inefficiency. See, for example, Aigner et al. (1977); Battese and Coelli (1988, 1992), and Greene (2005). It is very often assumed that the two-sided error is normally distributed and the one-sided error is truncated normally or exponentially. If so, the distribution of inefficiency conditional on the composed error is truncated normally. Given these conditional inefficiency distributions (one for each firm), a common empirical question is how does one assess relative inefficiency in the sample? There are essentially two approaches. The first approach is to calculate the mean of each conditional inefficiency distribution, using the value of the regression residual for each firm in the conditioning argument. See Jondrow et al. (1982) for the cross-sectional case and Battese and Coelli (1988) for the panel data case. These conditional means (evaluated at the residual values) can be ordered across firms, and a sample-wide view of inefficiency is inferred from the order statistic. In particular, the firm with the smallest conditional mean may be deemed efficient relative to the rest in the sample. A second approach is to use the conditional inefficiency distributions to calculate the probability that each firm is best (has the lowest inefficiency), conditional on the (joint) composed errors. See Horrace (2005). These conditional efficiency probabilities can be evaluated at the values of the (joint) regression residuals to provide an alternative view of (in)efficiency in the sample, and, in particular, the firm with the largest efficiency probability may be deemed the most efficient. The first approach is a marginal approach in that each conditional mean is derived from a single conditional inefficiency distribution. The second approach is simultaneous in that each conditional efficiency probability is derived from all the conditional distributions, jointly. In this sense, the conditional probabilities contain information from the efficiency rank statistic that the conditional means do not provide. In the parlance of the multiple comparisons and ranking and selection literatures (e.g., Bechhofer 1954; Dunnett 1955; Gupta 1956; Gupta 1955), the conditional efficiency probabilities account for the “multiplicity” in the rank statistic (e.g., firm 1 is better than firm 2 and firm 3 and…).

This paper extends the conditional probability statements of Horrace (2005) to calculate not only the conditional probability that each firm is best (lowest inefficiency), but also the conditional probabilities that each firm is any efficiency rank (best, 2nd best, …, 2nd worst, worst) in the sample. The suite of conditional probabilities provides a complete picture of efficiency in the sample and is informative. To see this, let the sample consist of n firms and let the unconditional distribution of efficiency be the same for each firm (a common assumption). Then, the unconditional probability that any firm is a particular efficiency rank is simply 1/n, an uninteresting result. That is, the unconditional probability of any particular efficiency rank can be characterized by a discrete uniform distribution across firms. Once we condition on the sample data (on the regression residuals), the shape of this distribution across firms becomes less uniform (more informative). It is in this sense that the proposed conditional efficiency probabilities are empirically useful. In fact, our simulations show that when the variance of the one-sided error is small relative to that of the two-sided error (a noisy experiment), the conditional probabilities are close to the unconditional result, 1/n. As noise decreases, the probability weights of being a particular efficiency rank shift across firms, so the distribution becomes more informative.
Given the suite of conditional efficiency rank probabilities (a partition of the event space that firm \( i \) is efficiency rank \( r \)), it is a simple matter to calculate the expected rank for each firm, conditional on the composed errors, evaluated at the residual values. These conditional expected ranks are also useful. Like the unconditional efficiency rank probabilities, the unconditional expected rank for each firm is constant across firms. For example, if \( n = 5 \) and if the unconditional distribution of inefficiency is (again) identical across firms, then the unconditional expected rank for each firm is \((1 + \cdots + 5)/5 = 3\), an uninteresting result. The conditional expected rank, however, varies across firms, and this variability informs our understanding of the efficiency rankings. Continuing the example, if the firm with the highest efficiency score has a conditional expected rank of 1.2 (1 being the best and 5 being the worst), we are much more confident that it is the best firm in the sample than if it has a conditional expected rank of 2.2, and the conditional expected rank of 1.2 is certainly more informative than its unconditional expected rank of 3. Not surprisingly, the informativeness of the conditional expected rank is increasing in the signal to noise ratio in our simulations. Continuing the example, the conditional expected rank of 1.2 for the firm with the highest inefficiency score might be from a less noisy experiment than the 2.2 result. In a very noisy experiment, the same conditional expected rank might be close to 3, the unconditional result. Our simulations also reveal interesting relationships between the skew of the one-sided error and the distribution of the conditional expected ranks across firms.

This paper is organized as follows. The next section presents the parametric frontier model, the conditional efficiency rank probabilities, and the conditional expected rank measure. The model allows for unbalanced panels, a case which has not been treated extensively in previous work on efficiency probabilities. In Sect. 3, a Monte Carlo study demonstrates how the empirical distribution of conditional efficiency rank probabilities and the conditional expected ranks vary with a) the amount of noise in the experiment and b) the skew of the unconditional inefficiency distributions. Section 4 presents an empirical application to vessel efficiency in the US North Atlantic Herring fleet, and Sect. 5 concludes.

2 Conditional inefficiency rank probabilities for parametric frontiers

We consider the parametric stochastic frontier model for an unbalanced panel of firms:

\[
y_{it} = \alpha + x_{it} \beta - u_i + v_{it}, \quad i = 1, \ldots, n, \quad t = 1, \ldots, T_i.
\]

(1)

Here, \( y_{it} \) are observed logarithm of output of the \( i \)th firm in the \( t \)th period, the \( x_{it} \) are observed production inputs, the \( u_i \geq 0 \) are \( iid \) unobserved errors representing unobserved inefficiency, and the \( v_{it} \) are \( iid \) unobserved errors that cause the efficiency frontier to be stochastic. We assume that the distribution of \( v_{it} \) is \( N(0, \sigma_v^2) \) and distribution of \( u_i \) is the truncation below zero of a \( N(\mu, \sigma_u^2) \) random variate.\(^1\) Other distributions for \( u_i \) have been considered (e.g., Greene 1990), but are beyond the

---

\(^1\) A fixed effect model is also considered in Schmidt and Sickles (1984).
scope of what follows. We also require that \( x_{it}, u_i, \) and \( v_{it} \) be independent. Since \( y_{it} \) are in log points, firm-level technical efficiency is defined as \( TE_i = \exp(-u_i) \). Maximum likelihood estimation of the model’s parameters \((\hat{\alpha}, \hat{\beta}, \hat{\mu}, \hat{\sigma_u^2} \) and \( \hat{\sigma_v^2} \) is consistent (as \( n \to \infty \) or as \( T_i \to \infty \)).

The model in (1) is fairly flexible. It can represent both Cobb–Douglas and trans-log specifications, and it can be recast as a cost, revenue, or profit function. Generalizations for time-varying \( u_i \) are plentiful. For example, see Kumbhakar (1990), Cornwell et al. (1990), Battese and Coelli (1992), Lee and Schmidt (1993), Cuesta (2000), Han and Orea (2005), Lee (2006) and Ahn and Lee (2007). There are also models that incorporate an additive firm heterogeneity term that is separate from inefficiency. See for example, Greene (2005), Chen et al. (2011), Kumbhakar et al. (2012), and Columbi et al. (forthcoming). Our empirical example in Sect. 4 involves a more flexible form than in (1), where the marginal products are allowed to vary across groups of firms; the model is estimated using the El-Gamal and Grether (1995, 2000) estimation classification algorithm. Regardless of the simplicity or complexity of the underlying production function specification, the proposed expected rank results that follow are applicable as long as the \textit{ex post} conditional distribution of technical inefficiency is truncated normal.²

Based on our assumptions in (1), Battese and Coelli (1988) show that the distribution of \( u_i \) conditional on the composed error \( \varepsilon_{it} = v_{it} - u_i \) is the truncation below zero of a \( N(\mu_{si}, \sigma_{si}^2) \) random variate with

\[
\mu_{si} = \frac{-\sigma_u^2 T_i \bar{\varepsilon}_i + \mu \sigma_v^2}{(T_i \sigma_u^2 + \sigma_v^2)}, \quad \bar{\varepsilon}_i = \frac{1}{T_i} \sum_{t=1}^{T_i} \varepsilon_{it} \quad \text{and} \quad \sigma_{si}^2 = \frac{\sigma_u^2 \sigma_v^2}{(T_i \sigma_u^2 + \sigma_v^2)}.
\]

That is, the conditional density function of \( u_i \) is

\[
f(u|\varepsilon_i) = \frac{1}{1 - \Phi(-\mu_{si}/\sigma_{si})} \exp \left\{ -\frac{(u - \mu_{si})^2}{2\sigma_{si}^2} \right\} \quad \text{with} \quad \varepsilon_i = [\varepsilon_{i1} \ldots \varepsilon_{iT_i}]'.
\]

Then the conditional distribution function is

\[
F(u|\varepsilon_i) = \int_{0}^{u} f(u|\varepsilon_i)du = \frac{\Phi([u - \mu_{si}]/\sigma_{si}) - \Phi(-\mu_{si}/\sigma_{si})}{1 - \Phi(\mu_{si}/\sigma_{si})},
\]

where \( \Phi \) is the cumulative distribution function of a standard normal random variate. Then the conditional mean of \( u_i \) is

² However, our results, the results of Flores-Lagunes et al. (2007) and the results of Horrace (2005) are easily generalizable to any parametric stochastic frontier model that produces a conditional inefficiency distribution of known form (based on assumptions on the error components).
\[ E(u|\varepsilon_i) = \mu_{si} + \sigma_{si} \frac{\phi(-\mu_{si}/\sigma_{si})}{\Phi(\mu_{si}/\sigma_{si})}, \]

with \( \phi \), the density of a standard normal random variate.

In principle, the population efficiency ranking is in terms of \( u_i \). That is, \( u_{[1]} \leq u_{[2]} \leq \cdots \leq u_{[n]} \), so that firm \([1]\) is most efficient in the population, and firm \([n]\) is least efficient. However, \( u_i \) is unobserved and cannot be directly estimated, so what is often done is to calculate the vector of residuals \( e_i = [e_{i1} \ldots e_{iT_i}]\) with \( e_{it} = y_{it} - \hat{\alpha} - x_{it} \hat{\beta} \), and estimate inefficiency as \( \hat{u}_i = E(u|\varepsilon_i = e_i) \), the conditional mean evaluated at \( \varepsilon_i = e_i \) (with \( \hat{\mu} = \mu, \sigma_u^2 = \hat{\sigma}_u^2 \), and \( \sigma_v^2 = \hat{\sigma}_v^2 \) for each firm. Empirical exercises often include a rank ordering of the \( \hat{u}_i \), which serves as a predictor of the ordered \( u_i \) conditional on \( \varepsilon_i \). If \( T_i = T \), so that \( \sigma_{si}^2 = \sigma_s^2 \), then the firm rankings based on \( \hat{u}_i \) will be identical to the rankings based on \( \hat{\varepsilon}_i \). However, in the case of unbalanced panels, it is possible that these two rankings will not be identical, because \( \sigma_{si}^2 \) causes \( \hat{u}_i \) to no longer be monotonic in \( \hat{\varepsilon}_i \).

Based on the distribution of \( u_i \) conditional on \( \varepsilon_i \), Horrace (2005) calculates the conditional probabilities that firm \( i \) is the most efficient in the sample, \( \Pr(i = [1]|\varepsilon_1, \ldots, \varepsilon_n) \), and least efficient in the sample \( \Pr(i = [n]|\varepsilon_1, \ldots, \varepsilon_n) \). The conditional efficiency probabilities are predicted by evaluating them at \( \varepsilon_i = e_i \), \( i = 1, \ldots, n \). We generalize those results to calculate the conditional probability that firm \( i \) is any efficiency rank, \( r \), in the sample, \( \Pr(i = [r]|\varepsilon_1, \ldots, \varepsilon_n) \). This conditional rank probability is the sum of the probabilities of all possible events where \( u_i \) is rank \( r \) (and there may be many), however it is not necessary to calculate all possible event permutations to determine it. Instead we can start from events in which a particular set of firms are more efficient than \( i \) and the remaining firms are less efficient, without regard for the rankings of the firms within each set. Consider one such event with firm \( i \) at rank \( r \), and define subsets \( N_{i}^{-}(r) = \{ j : u_j < u_i \} \) and \( N_{i}^{+}(r) = \{ j : u_j > u_i \} \), conditional on \( \varepsilon_1, \ldots, \varepsilon_n \). The conditional probability of this event is

\[
\int_{0}^{\infty} f(u|\varepsilon_i) \prod_{j \in N_{i}^{-} (r)} F(u|\varepsilon_j) \prod_{k \in N_{i}^{+}(r)} [1 - F(u|\varepsilon_k)] du.
\]

For any rank, except \( r = 1 \) or \( r = n \), there are multiple combinations of ranked firms above and below \( i \), which yield the same rank. In fact, there are \( n - 1 \) \( C_{r-1} = \frac{(n-1)!}{(n-r)! (r-1)!} \) combinations that have firm \( i \) at rank \( r \) out of \( n \) firms. Accordingly, we index the sets of firms above and below \( i \) for each different combination that produces the same rank as \( N_{i}^{-}(r) \) and \( N_{i}^{+}(r) \), \( l = 1, \ldots, n - 1 \) \( C_{r-1} \). Then the conditional efficiency rank probability for rank \( r \) out of \( n \) is

\[
\Pr(i = [r]|\varepsilon_1, \ldots, \varepsilon_n) = \sum_{l=1}^{n-1} C_{r-1} \int_{0}^{\infty} f(u|\varepsilon_i) \prod_{j \in N_{i}^{-} (r)} F(u|\varepsilon_j) \prod_{k \in N_{i}^{+}(r)} [1 - F(u|\varepsilon_k)] du,
\]

(2)
When \( r = 1 \) or \( r = n \), these reduce to the conditional efficiency probabilities of Horrace (2005). The \( n^2 \) probabilities in (2) can be predicted by evaluating them at \( \varepsilon_i = e_i, i = 1, \ldots, n \) (with \( \mu = \hat{\mu}, \sigma_u^2 = \hat{\sigma}_u^2 \), and \( \sigma_v^2 = \hat{\sigma}_v^2 \)). It is not difficult to generate computer algorithms for efficient calculation of these probabilities.\(^3\) When \( n \) is large, numerical calculation of the probabilities may be difficult, but simulating the probabilities by resampling from the conditional inefficiency distributions is straightforward. We employ both numerical and simulated probabilities in our empirical example.

Substituting the unconditional density function, \( f(u) \), and distribution function, \( F(u) \), for the conditional density function and distributions (respectively) in (2), it is clear that \( \Pr(i = [r]) = \Pr(j = [r]) \) with \( \sum_i \Pr(i = [r]) = 1 \), so that \( \Pr(i = [r]) = 1/n \). This argument hinges on the unconditional draws of \( u \) (over \( i \)) being identically distributed. Obviously, if the unconditional distribution of the \( u_i \) varies over \( i \) (e.g., Battese and Coelli 1995), then the unconditional \( \Pr(i = [r]) \) would not equal \( 1/n \) in general, and would be a function of the parameters of the underlying unconditional distributions. Also, if \( \sigma_v^2 \) is large relative to \( \sigma_u^2 \), then realizations of \( \varepsilon_i \) contain relatively little information about \( u_i \), so that the conditional distribution of \( u_i \) is close to its unconditional distribution. Therefore, when \( \sigma_v^2 \) is large relative to \( \sigma_u^2 \), the probabilities in (2) will be close to \( 1/n \).

Of course reporting the \( n^2 \) probabilities in (2) in an empirical exercise may be impractical. However, much of the pertinent information contained in the \( r = 1, \ldots, n \) conditional rank probabilities for a firm can be summarized with its conditional expected rank statistic,

\[
\rho_i = \sum_{r=1}^{n} r \Pr(i = [r] | \varepsilon_1, \ldots, \varepsilon_n) \in [1, n],
\]

\( i = 1, \ldots, n \). This measure is an alternative way to characterize efficiency ranks that accounts for multiplicity in the rank statistic through the probabilities in (2). Again, it can be predicted by evaluating \( \varepsilon_i \) at the values of \( e_i \) for every firm. It also responds to the relative magnitudes of the signal (\( \sigma_u^2 \)) and noise (\( \sigma_v^2 \)) in the same way as the probabilities in (2). In a particularly noisy setting, the conditional rank probabilities in (2) are approximately equal to \( 1/n \), and the conditional expected rank will be approximately equal across firms. In this sense, (2) and (3) provide information on one source of uncertainty in the efficiency ranks that the conditional means, \( E(u | \varepsilon_i) \), do not.\(^4\) Of course, the conditional mean, the conditional rank probabilities, and the conditional expected rank are all different measures, so comparisons of their abilities to serve as substitutes should not be overstated.

\(^3\) See, for example, Sect. 5.2.2 of Reingold et al. (1977) for an efficient algorithm that serves to find \( N_l^1 - (r) \) and \( N_l^1 + (r) \) for \( l = 1, \ldots, n-1 \) \( C_{r-1} \).

\(^4\) One could supplement the conditional means with the conditional prediction intervals of Horrace and Schmidt (1996), to judge how much the marginal distributions overlap. The degree of overlap may correspond to the extent to which the conditional probabilities and expected ranks are close to their unconditional counterparts, but this might be highly subjective and (perhaps) lead to an inaccurate assessment of the nature of efficiency in the population.
All of the different characterizations of inefficiency (and their relative rankings) are evaluated at \( \varepsilon_i = e_i \). Therefore, they all ignore estimation error, which (of course) is asymptotically negligible. Nonetheless, it may be important in finite samples. For the conditional means, there are existing procedures to address the issue. Simar and Wilson (2009) and Wheat et al. (forthcoming) recommend resampling techniques to incorporate estimation error into confidence intervals on technical inefficiency. Resampling techniques could certainly be employed to assess the effects of estimation error on the conditional expected ranks. The procedure to do so would be straightforward, and it would yield confidence sets for the vectors of conditional expected ranks. However, this is not the focus of the evaluation presented below.

3 Monte Carlo study

We use a series of simulations to demonstrate properties of the conditional expected rank statistic. For simplicity, we always set \( T_i = 1 \). As equation 2 shows, the conditional rank probabilities for each firm depend on the conditional distributions, \( f(u|\varepsilon_i) \), which themselves depend on three parameters \( \mu, \sigma_u^2, \) and \( \sigma_v^2 \). First, we follow standard simulation practice for stochastic frontier models (e.g., Olson et al. 1980), and explore how statistical noise, \( \sigma_v^2 \), affects the empirical distribution of the conditional rank probabilities in (2) and the conditional expected ranks in (3). To this end, we consider \( \sigma_v^2 = \{0.01, 0.1, 1, 10\} \) for fixed \( V(u) \). The point is that increasing noise should degrade the efficiency rank probabilities’ ability to accurately detect the true rank of any firm, so that the conditional expected ranks are increasingly uninformative. Second, Feng and Horrace (2012) show that the skew of the inefficiency distribution can also confound our detection of firm ranks at different ends of the order statistic in different ways.\(^5\) If the inefficiency distribution is “mostly stars” having many firms in the left tail \((u_i \cong 0 \text{ with high probability})\), then it is difficult to differentiate the individual ranks of these highly efficient firms (low \([r]\) firms). Conversely, if the inefficiency distribution is “mostly dogs” having fewer firms in the left tail \((u_i \cong 0 \text{ with low probability})\), then it is easier to differentiate the individual ranks of these highly efficient firms.\(^6\) The amount of relative mass in one tail of a distribution affects the skew of the distribution. Therefore, our second interest is in seeing the effects of distributional skew (for a fixed variance) on the conditional rank probabilities and the conditional expected ranks. To do this, we select values \( \mu \) and \( \sigma_u^2 \) that hold the variance constant at \( V(u) = 0.36 \) (the variance of a standard normal random variable truncated at zero) but produce skews of 0.5, 1.0, and 1.5, respectively.\(^7\) These values are listed in Table 1.

The different combinations of parameters \((\mu, \sigma_u^2, \sigma_v^2)\) yield a total of 12 separate exercises (four exercises for each of three skew levels). In each exercise, we use a total of 5,000 replications. We use a modest number of firms, \( n = 5 \), to reduce the computa-

---

\(^5\) Feng and Horrace are only concerned with detecting the best firm. We want to detect the rank of all firms.

\(^6\) The nomenclature “mostly stars and dogs” is due to Almanidis et al. (2014).

\(^7\) The skew of a truncated normal is necessarily positive. We use the “standardized” definition of skew where the 3rd central moment is divided by the third power of the standard deviation.
Table 1  Simulation parameters and the resulting truncated moments

| Underlying $N(\mu, \sigma_u^2)$ | Moments of the truncated distribution | Range of $(\sigma_v^2)$ |
|----------------------------------|--------------------------------------|-------------------------|
| Mean($\mu$) | Variance($\sigma_u^2$) | Mean | Variance | Skew | |
| 0.89 | 0.52 | 1.04 | 0.36 | 0.50 | 0.01–10 |
| 0.00 | 1.00 | 0.80 | 0.36 | 1.00 | 0.01–10 |
| −3.00 | 2.83 | 0.67 | 0.36 | 1.50 | 0.01–10 |

Fig. 1  Distribution of conditional rank probabilities with low skew: skew($u$) = 0.5

8 Again, the probabilities in (2) could be easily simulated for large $n$, but for the purposes of illustration, small $n$ is sufficient.
sample, firm 2 is the 2nd most efficient in the sample, ..., and firm 5 is the least efficient in the sample for any of our 5,000 Monte Carlo draws. These firm numbers, $i$, are along the x-axis, and the average conditional rank probabilities, $\Pr(i = [r] | \epsilon_1, ..., \epsilon_n)$, are along the y-axis on the graph. Consider $\Pr(i = [1] | \epsilon_1, ..., \epsilon_n)$, the dark blue series. Based on $\sigma^2_v = 0.01$, the (average conditional) probability that firm 1 is rank 1, $\Pr(i = [1] | \epsilon_1, ..., \epsilon_n)$, is 0.820; the probability that firm 2 is rank 1, $\Pr(i = [2] | \epsilon_1, ..., \epsilon_n)$, is 0.151; the probability that firm 3 is rank 1, $\Pr(i = [3] | \epsilon_1, ..., \epsilon_n)$, is 0.026; and $\Pr(i = [4] | \epsilon_1, ..., \epsilon_n) = \Pr(i = [5] | \epsilon_1, ..., \epsilon_n) = 0$. The probabilities for the other series are read from the graph similarly. For example, for $\sigma^2_v = 0.01$, we have $\Pr(i = [2] | \epsilon_1, ..., \epsilon_n) = 0.673$, $\Pr(i = [3] | \epsilon_1, ..., \epsilon_n) = 0.662$, $\Pr(i = [4] | \epsilon_1, ..., \epsilon_n) = 0.716$, and $\Pr(i = [5] | \epsilon_1, ..., \epsilon_n) = 0.867$. For the lowest noise experiment ($\sigma^2_v = 0.01$), we see that the analysis is better at correctly detecting firms with higher true $[r]$ than firms with lower true $[r]$ (compare $\Pr(i = [1] | \epsilon_1, ..., \epsilon_n) < \Pr(i = [5] | \epsilon_1, ..., \epsilon_n)$ and $\Pr(i = [2] | \epsilon_1, ..., \epsilon_n) < \Pr(i = [4] | \epsilon_1, ..., \epsilon_n)$), and this is always the case for our simulations (regardless of skew), because the distribution
of \( u \) will always have a thinner right tail (where \([r]\) is large) than left tail (where \([r]\) is small) as the skew of a truncated normal is always positive. However, the low skew (0.50) of the Fig. 1 simulations means that the distribution is relatively symmetric, so we shall see that differences in the ability of the conditional rank probabilities to accurately detect high- and low-ranked firms will become even more stark as we increase the skew (and increase uncertainly over which firms have lower true \([r]\)). See Figs. 4, 5, and 6 for a typical empirical inefficiency distribution for each of our three levels of skew: 0.5, 1.0, and 1.5, respectively. Each figure is a kernel density plot using a Gaussian kernel, bandwidth chosen by Silverman’s rule of thumb (Silverman 1986), and no boundary-bias correction.

Continuing with the low-skew results of Fig. 1, as we increase \( \sigma_v^2 = \{0.01, 0.1, 1, 10\} \), the empirical distribution of the efficiency probabilities becomes more uniform (and less informative). However, we also see in the four panels of Fig. 1 that it is always the case that \( \Pr(1 = [1]|\varepsilon_1, \ldots, \varepsilon_n) < \Pr(5 = [5]|\varepsilon_1, \ldots, \varepsilon_n) \), even in the noisiest (\( \sigma_v^2 = 10 \)) panel. Both of these empirical phenomena remain as we increase the skew (asymmetry) of the distribution of \( u \) to 1.0 and to 1.5 in Figs. 2 and 3, respectively (while holding \( V(u) \) constant). Looking across the figures, we see the effect. Consider the lowest noise panel (upper-left panel) in Figs. 1, 2, and 3. As the skew increases across Figs. 1, 2, and 3, \( \Pr(1 = [1]|\varepsilon_1, \ldots, \varepsilon_n) \) is decreasing (0.820, 0.754, and 0.699, respectively), while \( \Pr(5 = [5]|\varepsilon_1, \ldots, \varepsilon_n) \) is slightly increasing (0.867, 0.873, and 0.879, respectively). In the words of Almanidis et al. (2014), when the conditional distribution of \( u \) has “fewer stars” (low skew of Fig. 4), it is easier to detect stars, \( \Pr(1 = [1]|\varepsilon_1, \ldots, \varepsilon_n) = 0.820 \), than when there are “mostly stars” (high skew of Fig. 6), \( \Pr(1 = [1]|\varepsilon_1, \ldots, \varepsilon_n) = 0.696 \).9 These are inferential insights

---

9 These results are consistent with the Feng and Horrace (2012).
that the conditional means, $E(u|\varepsilon)$, would not uncover. These are also manifest in the conditional expected ranks which we now consider.

Once the conditional rank probabilities are calculated for each firm at each rank, calculation of the conditional expected ranks of equation 3 is straight-forward. The distributions of conditional expected rank (for each simulation run in Figs. 1, 2, and 3) are contained in Tables 2, 3, and 4 (respectively). The utility of the conditional expected ranks is immediately obvious. First, the extent to which noise affects $\Pr(i = [r]|\varepsilon_1, \ldots, \varepsilon_n)$ is clear. Consider the first panel ($\sigma^2_v = 0.01$) of Table 2. The difference
between the true rank of firm 1 (first column) and the average conditional expected rank (second column) is relatively small \((1 - 1.21 = -0.21)\), but this difference is increasing in magnitude as we read down the panels and the level of noise increases: \(1 - 1.72 = -0.72, 1 - 2.66 = -1.66, \) and \(1 - 2.96 = -1.96\). These qualitative results are true for all firms (true rank) and for all levels of skew (Tables 2, 3 and 4). Obviously, as noise increases, the conditional expected ranks are moving toward the unconditional expected rank, 3 (bottom panel in Table 2), which reflects the nearly uniform distribution of the conditional efficiency probabilities (bottom panel of Fig. 1).

The response of the quantiles of the expected ranks (columns with the heading “Quantiles of \(\rho\)”) to increased noise is clear in Tables 2, 3, and 4: noise tends to push extreme quantiles (and their surrounding probability mass) to the center of the empirical distribution of the conditional expected ranks. Also, the effect of skew is clear across the tables. Consider firms 1 and 5 in the first (low noise) panels of Tables 2, 3, and 4. For firm 1, the difference between its true rank [1] and the average conditional expected rank is increasing in magnitude \((-0.21, -0.30, -0.39)\) as skew increases.
(0.5, 1.0, and 1.5) across Tables 2, 3, and 4, respectively, while the same differences for firm 5 are non-increasing in magnitude across the tables (0.16, 0.15, 0.15). Again, this reflects the fact that as skew increases (and there are relatively more stars in the inefficiency distribution) it is harder to detect “stars” in the left tail of the inefficiency distribution than to detect “dogs” in the right tail. The quantiles in Tables 2, 3, and 4 also seem to nest important information about the conditional efficiency ranks. In the first panel of Table 2, the rank of firm 1 is correctly estimated by the 10 and 25 % quantiles, those of firm 2, 3, and 4 by the 50 % quantile, while the efficiency rank of firm 5 is estimated precisely by the 75 and 90 % quantiles. Moving to the third panel from the top, we can find that the efficiency rank is estimated best by the 10 % quantile for firm 1, the 50 % quantile for firm 3, and the 90 % quantile for firm 5.\footnote{We thank an anonymous referee for these insights into the information contained in the quantiles, which may imply that a resampling method, such as bootstrap, may provide better estimates of expected rank.}

\begin{table}[h]
\centering
\caption{Mean and quantiles of conditional expected rank, medium skew (skew(u) = 1.0)}
\begin{tabular}{cccccc}
\hline
\text{True rank} & \text{Mean } \rho & \text{Quantiles of } \rho \\
 & & & \text{10 %} & \text{25 %} & \text{50 %} & \text{75 %} & \text{90 %} \\
\hline
\text{\(\sigma^2_v = 0.01\)} & & & & & & & \\
1 & 1.30 & 1.00 & 1.01 & 1.14 & 1.48 & 1.83 \\
2 & 2.04 & 1.44 & 1.75 & 2.00 & 2.26 & 2.67 \\
3 & 2.94 & 2.26 & 2.70 & 2.99 & 3.17 & 3.58 \\
4 & 3.87 & 3.26 & 3.73 & 3.99 & 4.03 & 4.34 \\
5 & 4.85 & 4.43 & 4.85 & 5.00 & 5.00 & 5.00 \\
\text{\(\sigma^2_v = 0.1\)} & & & & & & & \\
1 & 1.86 & 1.20 & 1.41 & 1.74 & 2.19 & 2.65 \\
2 & 2.30 & 1.53 & 1.81 & 2.18 & 2.72 & 3.21 \\
3 & 2.86 & 1.94 & 2.34 & 2.83 & 3.34 & 3.83 \\
4 & 3.55 & 2.50 & 3.06 & 3.65 & 4.07 & 4.47 \\
5 & 4.43 & 3.48 & 4.13 & 4.67 & 4.93 & 4.99 \\
\text{\(\sigma^2_v = 1\)} & & & & & & & \\
1 & 2.65 & 2.01 & 2.25 & 2.60 & 2.98 & 3.34 \\
2 & 2.78 & 2.12 & 2.39 & 2.73 & 3.14 & 3.52 \\
3 & 2.95 & 2.26 & 2.53 & 2.90 & 3.32 & 3.70 \\
4 & 3.14 & 2.39 & 2.70 & 3.12 & 3.55 & 3.94 \\
5 & 3.48 & 2.65 & 3.03 & 3.49 & 3.95 & 4.33 \\
\text{\(\sigma^2_v = 10\)} & & & & & & & \\
1 & 2.95 & 2.67 & 2.79 & 2.95 & 3.10 & 3.24 \\
2 & 2.97 & 2.69 & 2.82 & 2.97 & 3.12 & 3.26 \\
3 & 3.00 & 2.72 & 2.84 & 2.99 & 3.14 & 3.29 \\
4 & 3.02 & 2.73 & 2.85 & 3.01 & 3.17 & 3.32 \\
5 & 3.06 & 2.76 & 2.90 & 3.06 & 3.22 & 3.36 \\
\hline
\end{tabular}
\end{table}
Table 4  Mean and quantiles of conditional expected rank, high skew (skew(u) = 1.5)

| True rank | Mean ρ | Quantiles of ρ |
|-----------|--------|----------------|
|           |        | 10 % | 25 % | 50 % | 75 % | 90 % |
| σ²_v = 0.01 |       |      |      |      |      |      |
| 1         | 1.39  | 1.00 | 1.04 | 1.26 | 1.60 | 1.95 |
| 2         | 2.04  | 1.43 | 1.70 | 2.00 | 2.30 | 2.73 |
| 3         | 2.88  | 2.14 | 2.58 | 2.97 | 3.14 | 3.55 |
| 4         | 3.84  | 3.20 | 3.67 | 3.98 | 4.01 | 4.31 |
| 5         | 4.85  | 4.46 | 4.87 | 5.00 | 5.00 | 5.00 |
| σ²_v = 0.1 |       |      |      |      |      |      |
| 1         | 1.96  | 1.27 | 1.50 | 1.85 | 2.32 | 2.80 |
| 2         | 2.32  | 1.54 | 1.82 | 2.22 | 2.71 | 3.25 |
| 3         | 2.79  | 1.90 | 2.26 | 2.75 | 3.27 | 3.76 |
| 4         | 3.48  | 2.44 | 2.95 | 3.56 | 4.03 | 4.42 |
| 5         | 4.44  | 3.45 | 4.13 | 4.69 | 4.95 | 5.00 |
| σ²_v = 1  |       |      |      |      |      |      |
| 1         | 2.63  | 1.90 | 2.17 | 2.55 | 3.00 | 3.48 |
| 2         | 2.76  | 2.00 | 2.29 | 2.70 | 3.18 | 3.62 |
| 3         | 2.90  | 2.12 | 2.41 | 2.85 | 3.33 | 3.82 |
| 4         | 3.14  | 2.29 | 2.63 | 3.11 | 3.62 | 4.06 |
| 5         | 3.57  | 2.60 | 3.03 | 3.58 | 4.13 | 4.55 |
| σ²_v = 10 |       |      |      |      |      |      |
| 1         | 2.93  | 2.51 | 2.69 | 2.91 | 3.15 | 3.39 |
| 2         | 2.96  | 2.52 | 2.72 | 2.94 | 3.19 | 3.42 |
| 3         | 2.98  | 2.56 | 2.73 | 2.97 | 3.21 | 3.44 |
| 4         | 3.02  | 2.59 | 2.78 | 3.01 | 3.26 | 3.48 |
| 5         | 3.10  | 2.64 | 2.85 | 3.09 | 3.34 | 3.58 |

Finally, the conditional expected ranks are a convenient normalization of relative efficiency. Notice that the normalization is pegged to both ends of the true order statistic (1 and 5), such that \( \rho_i \in [1, n] \). Compare this to the traditional predictor of \( TE_i = \exp(-u_i) \),

\[
\bar{TE}_i = E\left[ \exp(-u) | \varepsilon_i \right] = \frac{1 - \Phi\left[ \frac{\sigma \left( \frac{\mu_{ni}}{\sigma} \right)}{\sigma} \right]}{1 - \Phi\left[ -\frac{\mu_{ni}}{\sigma} \right]} \exp\left\{ -\mu_{ni} + \frac{1}{2} \sigma^2 \right\},
\]
evaluated at \( \varepsilon_i = e_i \). (See Jondrow et al. 1982.) This absolute predictor normalizes efficiency predictions to the unit interval, \( \bar{TE}_i \in (0, 1) \). Therefore, linear renormalizations of expected rank, like \( 1 - (1 - \rho_i) / n \), can be thought of as alternatives to the \( \bar{TE}_i \) normalization. However, the former is measured on a relative (within sample) scale, while the latter is measured on an absolute (out of sample) scale.
4 Empirical example

To illustrate our results on expected ranks, we revisit the empirical exercise in Flores-Lagunes et al. (2007), who estimate a stochastic production frontier for an unbalanced panel for $n = 39$ vessels from the US North Atlantic Herring fleet (2000-2003). They specify a heterogeneous production function and use the El-Gamal and Grether estimation classification algorithm (El-Gamal and Grether 1995, 2000) to classify the fleet into three production tiers. See Flores-Lagunes et al. (2007) for a complete discussion of the data, the production function and the estimation algorithm. Suffice it to say that vessel output is total catch (tons) and inputs are things like vessel size (tons), hours at sea, and crew size. The estimation yields $\mu_{ni}$ and $\sigma^2_{ni}$ for each vessel. That is, each vessel’s conditional inefficiency distribution is a $N(\mu_{ni},\sigma^2_{ni})$ truncated at zero.

The North Atlantic Herring fleet consists of two technologies: trawlers and “purse seiners”. While in motion, trawling vessels drag large nets to take catch. A purse seine is a large net that is dropped toward the ocean floor, while the vessel is at rest. The gear encircles catch as it is hauled back up to the boat. Vessels use only one of these technology (there are costs to refitting vessels with the different gear types). The El-Gamal and Grether estimation classification algorithm stratifies the fleet into three production tiers, where each tier has separate marginal product estimates (estimates of $\alpha$ and $\beta$ in Eq. 1). The first and second tiers consist exclusively of trawlers and the third tier consists of a mix of trawlers and purse seiners. Efficiency is characterized within (and not across) each production tier. Flores-Lagunes et al. (2007) also calculate 95% subsets of the best and worst vessels in each tier. That is, they calculate subsets of vessel indices which contain the index of the most and least efficient vessels in the population with probability at least 95%. We use their subsets of the best in each tier in what follows.

The estimates of $\mu_{ni}$ and $\sigma^2_{ni}$ for the all vessels in each production tier (tier 1, tier 2, and tier 3) are reproduced in the second and third columns of Tables 5, 6, and 7, respectively. The first column contains the unique vessel numbers from the Flores-Lagunes et al. (2007) analysis. The fourth column contains traditional technical efficiency predictors $\hat{\text{TE}}_i = E[\exp(-u) | \varepsilon_i]$ from (4) evaluated at $\varepsilon_i = e_i$, and the results in Tables 5, 6, and 7 are ranked on this value. The fifth column contains the conditional expected ranks, $\rho_i$, in (3) for the five vessels with the largest $\hat{\text{TE}}$ in each tier. The choice of five vessels is purely for convenience, as numerical calculation of the rank probabilities becomes difficult with more vessels. The sixth column also contains conditional expected ranks, but for only those vessels that were in the subset of the best vessels with 95% probability. For example, in tier 1 (Table 5), the 95% subset of the best is $\{1, 2, 5, 10, 11, 14, 27, 39, 37\}$, and we calculate conditional expected ranks for these vessels only. For the expected ranks in column 6, we used simulated conditional rank probabilities, because the subsets were often larger than

---

11 The results of the estimation are not reproduced here to focus attention on the different characterization of efficiency ranks and the importance of the proposed conditional expected rank statistic.

12 The expected ranks are calculated using the conditional efficiency rank probabilities for only these five vessels in each tier. In particular we did not calculate the efficiency rank probabilities for all vessels, and calculate expected rank only for these five, based on the probabilities from all vessels.
Table 5 Heterogeneous vessel efficiency results tier 1, sorted on $\hat{TE}_i$

| Vessel $i$ | $\sigma_i^2$ | $\mu_i^*$ | $\hat{TE}_i$ | $\rho_i$ Top 5<sup>a</sup> | $\rho_i$ Subset<sup>b</sup> | $\rho_i$ All<sup>c</sup> |
|-----------|--------------|------------|--------------|----------------|----------------|----------------|
| 2         | 0.001        | 0.074      | 0.928        | 2.272          | 3.181          | 3.360          |
| 39        | 0.013        | 0.03       | 0.905        | 2.698          | 3.829          | 4.165          |
| 11        | 0.011        | 0.094      | 0.881        | 3.157          | 4.576          | 5.026          |
| 5         | 0.032        | 0.04       | 0.859        | 3.394          | 5.031          | 5.584          |
| 14        | 0.064        | -0.077     | 0.846        | 3.479          | 5.176          | 5.846          |
| 27        | 0.345        | -1.553     | 0.844        | -             | 5.092          | 5.821          |
| 10        | 0.125        | -0.360     | 0.841        | -             | 5.192          | 5.962          |
| 37        | 0.039        | 0.126      | 0.817        | -             | 5.966          | 6.829          |
| 4         | 0.020        | 0.242      | 0.781        | -             | -              | 7.958          |
| 1         | 0.183        | 0.034      | 0.725        | -             | 6.950          | 8.498          |
| 17        | 0.039        | 0.303      | 0.731        | -             | -              | 8.982          |
| 28        | 0.026        | 0.387      | 0.685        | -             | -              | 9.977          |
| 38        | 0.039        | 1.230      | 0.298        | -             | -              | 12.989         |

<sup>a</sup> Expected rank among only the top 5 vessels sorted on $\hat{TE}_i$

<sup>b</sup> Expected rank among only the vessels in subset \{1, 2, 5, 10, 11, 14, 27, 39, 37\}

<sup>c</sup> Expected rank among all vessels in the tier

Table 6 Heterogeneous vessel efficiency results tier 2, sorted on $\hat{TE}_i$

| Vessel $i$ | $\sigma_i^2$ | $\mu_i^*$ | $\hat{TE}_i$ | $\rho_i$ Top 5<sup>a</sup> | $\rho_i$ Subset<sup>b</sup> | $\rho_i$ All<sup>c</sup> |
|-----------|--------------|------------|--------------|----------------|----------------|----------------|
| 21        | 0.011        | 0.041      | 0.907        | 2.535          | 3.129          | 3.298          |
| 19        | 0.039        | -0.182     | 0.903        | 2.547          | 3.204          | 3.200          |
| 13        | 0.009        | 0.135      | 0.864        | 3.318          | 4.180          | 4.345          |
| 12        | 0.125        | -0.518     | 0.863        | 3.067          | 3.954          | 4.143          |
| 7         | 0.345        | -1.129     | 0.814        | 3.533          | 4.718          | 5.046          |
| 32        | 0.095        | 0.095      | 0.768        | -             | 5.493          | 5.952          |
| 6         | 0.095        | 0.126      | 0.758        | -             | 5.669          | 6.099          |
| 8         | 0.183        | -0.047     | 0.743        | -             | 5.651          | 6.239          |
| 18        | 0.055        | 0.332      | 0.705        | -             | -              | 7.028          |
| 15        | 0.001        | 0.692      | 0.501        | -             | -              | 9.716          |
| 24        | 0.009        | 0.967      | 0.382        | -             | -              | 11.217         |
| 9         | 0.002        | 1.032      | 0.357        | -             | -              | 11.712         |

<sup>a</sup> Expected rank among only the top 5 vessels sorted on $\hat{TE}_i$

<sup>b</sup> Expected rank among only the vessels in subset \{6, 7, 8, 12, 13, 19, 21, 32\}

<sup>c</sup> Expected rank among all vessels in the tier

5, making numerical calculation difficult. Finally, the last column in Tables 5, 6 and 7 contains the conditional expected ranks for all vessels in each tier, also based on simulated rank probabilities.<sup>13</sup>

<sup>13</sup> When conditional expected ranks are based on simulated probabilities, the simulation sample size is always 5,000.
Table 7  Heterogeneous vessel efficiency results tier 3, sorted on $\hat{TE}_i$

| Vessel | $\sigma_i^2$ | $\mu_i^*$ | $\hat{TE}_i$ | $\rho_i$ Top 5<sup>a</sup> | $\rho_i$ Subset<sup>b</sup> | $\rho_i$ All<sup>c</sup> |
|--------|-------------|-----------|-------------|----------------|----------------|----------------|
| 3      | 0.001       | 0.001     | 0.971       | 1.395          | 1.387          | 1.442          |
| 33     | 0.002       | 0.065     | 0.934       | 2.528          | 2.524          | 2.738          |
| 16     | 0.008       | 0.011     | 0.930       | 2.500          | 2.456          | 2.753          |
| 30     | 0.001       | 0.193     | 0.825       | 4.425          | –              | 5.314          |
| 34     | 0.023       | 0.173     | 0.817       | 4.151          | 3.631          | 5.189          |
| 35     | 0.006       | 0.215     | 0.808       | –              | –              | 5.568          |
| 20     | 0.048       | 0.179     | 0.783       | –              | –              | 5.806          |
| 31     | 0.001       | 0.494     | 0.611       | –              | –              | 8.783          |
| 29     | 0.007       | 0.504     | 0.606       | –              | –              | 8.993          |
| 23     | 0.125       | 0.604     | 0.555       | –              | –              | 9.836          |
| 26     | 0.006       | 0.602     | 0.550       | –              | –              | 10.137         |
| 22     | 0.006       | 0.836     | 0.435       | –              | –              | 11.980         |
| 36     | 0.010       | 0.913     | 0.403       | –              | –              | 12.511         |
| 25     | 0.006       | 1.230     | 0.293       | –              | –              | 14.263         |

<sup>a</sup> Expected rank among only the top 5 vessels sorted on $\hat{TE}_i$

<sup>b</sup> Expected rank among only the vessels in subset {3, 16, 33, 34}

<sup>c</sup> Expected rank among all vessels in the tier

The results are compelling. Starting with Table 5, we see that for the top five vessels (column 5) the conditional expected ranks only range in value from 2.272 (vessel 2) to 3.479 (vessel 14), indicating a fairly noisy analysis. Had this been a particularly precise empirical exercise the lower bound of the range would be closer to 1 and the upper bound closer to 5. This is echoed in the subset of the nine best vessels (column 6) where the conditional expected ranks only range in value from 3.181 (vessel 2) to 6.950 (vessel 1), and in the full complement of thirteen vessels in the tier (column 7) where it ranges from 3.360 to 12.989. As we increase the number of vessels included in the conditional expected ranks across columns 5 through 7, we see that the ability of the conditional expected ranks to cover the population range of vessel ranks is diminished. For example, the expected rank for the top-ranked vessel, number 2, is increasing from 2.272 to 3.181 to 3.360. This is presumably because of the multiplicity in the conditional efficiency rank probability calculation in Eq. 2. Also the vessel rankings based on $\hat{TE}_i$ always match those based on $\rho_i$ for the top five vessels in column 5 of Table 5. This is almost the case in columns 6 and 7, except for vessels 14 and 27 whose expected ranks are reversed compared to the $\hat{TE}_i$ ranks. However, these are small (perhaps, negligible) differences which could be due to sampling variability in the simulated rank probabilities.

Table 6 tells a slightly different story. The range of the conditional expected ranks is tighter than in Table 5 and only ranges from 2.535 (vessel 21) to 3.533 (vessel 7) for the top 5 vessels (column 5), so the analysis is more noisy, however, the ranks based on $\hat{TE}_i$ are different than those based on $\rho_i$ for the top five. In particular, the ranks of vessel 13 and 12 are reversed, and it is clear why this is the case: the truncated normal distributions (upon which they are based) are vastly different in shape even...
though the means are approximately the same. That is, $E[\exp(-u)|\varepsilon_{12} = \epsilon_{12}] = 0.863 \cong E[\exp(-u)|\varepsilon_{13} = \epsilon_{13}] = 0.864$, however, the means and variances of the distributions before truncation are extremely different. Compare $\mu_{*13} = 0.135$ to $\mu_{*12} = -0.518$ and $\sigma_{*13}^2 = 0.009$ to $\sigma_{*12}^2 = 0.125$. Vessel 12 has more mass near zero in the distribution of $f(u|\varepsilon_i)$. This difference in the ranks between $\hat{TE}_i$ and $\rho_i$ for vessels 12 and 13 persists as we move to the subset of the eight best vessels (column 6) and all vessels (column 7).

Table 7 tells an even more nuanced story. The range of the expected ranks for the top 5 vessels (column 5) is wider: from 1.395 (vessel 3) to 4.151 (vessel 34), so this is the most precise rank statistic across the same columns in Tables 5, 6 and 7, yet there is still some switching in the ranks based on $\rho_i$ for these five vessels. In particular, the ranks of vessels 33 and 16 and of 30 and 34 are switched in column 5. Notice that the differences in $\hat{TE}_i$ for these vessel pairs are not that large, so it is really not surprising that the additional information provided by the conditional rank probabilities might switch the ranking. They are also switched in column 6 for vessels 33 and 16. Notice that in column 6, vessel 30 did not make it into the subset of the best vessels, while vessel 34 did. This is because the variance of inefficiency (before truncation) is much smaller for vessel 30 ($\sigma_{*30}^2 = 0.001$) than it is for vessel 34 ($\sigma_{*34}^2 = 0.023$). That is, we cannot reject the hypothesis that 34 is best, while we can for vessel 30. (A similar phenomenon occurs for vessels 4 and 1 in column 6 of Table 5). All of this underscores the importance of taking into account multiplicity and noise in any ranking exercise.

### 5 Conclusions

We extend the current literature on ranked efficiency scores by defining and proposing the use of conditional efficiency rank probabilities and conditional expected efficiency ranks as a means to provide improved insight into efficiency score rankings. Although our model was fairly restrictive, our results can be more broadly applied than this might indicate. Indeed, there is a broad class of parametric models that yield conditional efficiency distributions that are truncated normal and to which our results directly apply. Additionally, even if the resulting conditional distributions are not truncated normals, our results (and the results of Horrace 2005) can be adapted to these cases.

We demonstrated nuances of the proposed measures with a Monte Carlo Study. The conditional expected ranks responded in predictable ways to the inherent noisiness of the statistical exercise and to the skew of the underlying efficiency distribution. While it is generally ignored in empirical applications of the stochastic frontier model, skew is a very important moment to consider in drawing conclusions on ranked efficiency predictors. Our empirical example based on fishing vessels underscores the importance of taking into account multiplicity and noise in any ranking exercise, and the empirical relevance of the conditional rank probabilities and the conditional expected ranks is made clear.

One potential area of future research is that the OLS residuals are necessarily correlated, so while the conditional inefficiency distributions based on the true regression errors are independent, these distributions based on the residuals are technically not so. It would be interesting to see if analytic solutions were forthcoming and the corre-
lation of the residual could be estimated or approximated. It may also be worthwhile to consider resampling techniques to estimate confidence intervals for the conditional expected ranks, so that the usual assumption \( \hat{\beta} = \beta \) can be relaxed.

It may be fruitful to explore higher moments of the conditional rank distribution for each firm. We have discussed the conditional expectation of the distribution, but it may be worthwhile to consider the conditional variance of the rank of each firm. Calculating the variance, and any higher moments, would be a straight-forward exercise based on the conditional rank probabilities that we have presented. One might speculate that firms with high conditional probabilities of being best and worst would have higher conditional variance of their rank distribution than those with high probability of being in the center of the efficiency rank statistic. The best and worst firms will have more weight in one tail of their conditional rank distributions than firms with higher probability at the median efficiency ranks. However, this remains to be seen.

Acknowledgments The authors are grateful to Peng Liu for excellent research assistance and to several anonymous referees for their suggestions.

References

Ahn SC, Lee YH (2007) Panel data models with multiple time-varying individual effects. J Prod Anal 27:1–12
Aigner DJ, Lovell CAK, Schmidt P (1977) Formulation and estimation of stochastic frontier production functions. J Econom 6:21–37
Almanidis P, Qian J, Sickles RC (2014) Stochastic frontiers with bounded inefficiency. In: Sickles RC, Horrace WC (eds) Festschrift in honor of Peter Schmidt: Econometric methods and applications. Springer Science & Business Media, New York, pp 47–81
Battese GE, Coelli TJ (1988) Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. J Econom 38:387–399
Battese GE, Coelli TJ (1992) Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India. J Prod Anal 3:153–170
Battese GE, Coelli TJ (1995) A model for technical inefficiency effects in a stochastic frontier production function for panel data. Empir Econ 20:325–332
Bechhofer RE (1954) A single-sample multiple decision procedure for ranking means of normal populations with known variances. Ann Math Stat 25:16–39
Chen YY, Schmidt P, Wang HJ, (2011) Consistent estimation of the fixed effects stochastic frontier model. Unpublished manuscript
Colombi R, Kumbhakar SC, Martinix G, Vittadini G (forthcoming) Closed-skew normality in stochastic frontiers with individual effects and long/short-run efficiency. J Prod Anal
Cornwell C, Schmidt P, Sickles R (1990) Production frontiers with cross-sectional and time-series variation in efficiency levels. J Econom 46:185–200
Cuesta RA (2000) A production model with firm-specific temporal variation in technical efficiency: with application to Spanish dairy farms. J Prod Anal 13:139–158
Dunnett CW (1955) A multiple comparison procedure for comparing several treatments with a control. J Am Stat Assoc 50:1096–1121
El-Gamal M, Grether D (1995) Are people Bayesian? Uncovering behavior strategies. J Am Stat Assoc 90:1137–1145
El-Gamal M, Grether D (2000) Changing decision rules: Uncovering behavioral strategies using estimation classification (EC). In: Machina M et al. (eds) Preferences, beliefs, and attributes in decision making. Kluwer, New York
Feng Q, Horrace WC (2012) Alternative technical efficiency measures: skew, bias and scale. J Appl Econom 27:253–268
Flores-Lagunes A, Horrace WC, Schnier KE (2007) Identifying technically efficient fishing vessels: a non-empty, minimal subset approach. J Appl Econom 22:729–745
Greene WH (1990) A gamma distributed stochastic frontier model. J Econom 46:141–163
Greene WH (2005) Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. J Econom 126:269–303
Gupta SS (1956) On a decision rule for a problem of ranking means. Institute of Statistics Mimeo Series No. 150, University of North Carolina
Gupta SS (1965) On some multiple decision (selection and ranking) rules. Technometrics 7:225–245
Han C, Orea L (2005) Estimation of panel data model with parametric temporal variation in individual effects. J Econom 126:241–267
Horrace WC (2005) On ranking and selection from independent truncated normal distributions. J Econom 126:335–354
Horrace WC, Schmidt P (1996) Confidence statements for efficiency estimates from stochastic frontier models. J Prod Anal 7:257–282
Jondrow J, Lovell CAK, Materov IS, Schmidt P (1982) On the estimation of technical efficiency in the stochastic production function model. J Econom 19:233–238
Kumbhakar SC (1990) Production frontiers, panel data, and time-varying technical inefficiency. J Econom 46:201–211
Kumbhakar SC, Lien G, Hardaker JB (2012) Technical efficiency in competing panel data models: a study of Norwegian grain farming. J Prod Anal. doi: 10.1007/s11123-012-0303-1
Lee YH (2006) A stochastic production frontier model with group-specific temporal variation in technical efficiency. Eur J Oper Res 174(3):1616–1630
Lee YH, Schmidt P (1993) A production frontier model with flexible temporal variation in technical inefficiency. In: Fried H, Lovell CAK, Schmidt P (eds) The measurement of productivity efficiency: techniques and applications. Oxford University Press, Oxford
Olson JA, Schmidt P, Waldman DM (1980) A Monte Carlo study of estimators of stochastic frontier production functions. J Econom 13:67–82
Reingold EM, Nievergelt J, Deo N (1977) Combinatorial algorithms: theory and practice. Prentice Hall, New York
Schmidt P, Sickles RC (1984) Production frontiers and panel data. J Bus Econ Stat 2:367–374
Silverman BW (1986) Density estimation for statistics and data analysis. Chapman and Hall, London
Simar L, Wilson PW (2009) Inferences from cross-sectional, stochastic frontier models. Econom Rev 29:62–98
Wheat P, Greene WH, Smith A (forthcoming) Understanding prediction intervals for firm specific inefficiency scores from parametric stochastic frontier models. J Prod Anal