

IDENTIFYING TECHNICALLY EFFICIENT FISHING VESSELS: A NON-EMPTY, MINIMAL SUBSET APPROACH

ALFONSO FLORES-LAGUNES,^a WILLIAM C. HORRACE^{b*} AND KURT E. SCHNIER^c

^a *Department of Economics, University of Arizona, Tucson, Arizona, USA*

^b *Center for Policy Research, Syracuse University, Syracuse, New York, USA*

^c *Department of Environmental and Natural Resource Economics, University of Rhode Island, Kingston, Rhode Island, USA*

SUMMARY

Stochastic frontier models are often employed to estimate fishing vessel technical efficiency. Under certain assumptions, these models yield efficiency measures that are means of truncated normal distributions. We argue that these measures are flawed, and use the results of Horrace (2005) to estimate efficiency for 39 vessels in the Northeast Atlantic herring fleet, based on each vessel's probability of being efficient. We develop a subset selection technique to identify groups of efficient vessels at pre-specified probability levels. When homogeneous production is assumed, inferential inconsistencies exist between our methods and the methods of ranking the means of the technical inefficiency distributions for each vessel. When production is allowed to be heterogeneous, these inconsistencies are mitigated. Copyright © 2007 John Wiley & Sons, Ltd.

Received 4 April 2005; Revised 4 May 2006

1. INTRODUCTION

There is a growing resource economics literature concerned with the estimation of the technical efficiency of fishing vessels utilizing the stochastic frontier model. In this setting, the stochastic frontier model is motivated for two reasons. The first is the need to address the excess capacity prevalent in many of our fisheries, and the second centers on estimating the determinants of 'skipper skill' to determine the characteristics of 'highliners': vessel captains who consistently outfish others in the fleet. Excess capacity is conventionally estimated by comparing the current short-run production possessed by a vessel to that which could be achieved at the maximum utilization of their production inputs (Felthoven, 2002; Kirkley *et al.*, 2002, 2004). Estimation of 'skipper skill' is achieved by regressing vessel technical efficiencies on captain demographic variables, and it is often concluded that 'skipper skill' is positively correlated with a captain's experience and education (Kirkley *et al.*, 1998; Sharma and Leung, 1998; Pascoe and Coglán, 2002). Given the twofold importance of the stochastic frontier model in the fisheries economics literature, the inferential properties of technical efficiency measures need to be further investigated to assess their reliability in advising policy. This is the purpose of this research effort.

In this context the parametric stochastic frontier model (Meeusen and van den Broeck, 1977; Aigner *et al.*, 1977) yields estimates of the parameters (mean and variance) of a normal distribution for each vessel, which (when truncated at zero) represents the distribution of a vessel's inefficiency

* Correspondence to: William C. Horrace, Center for Policy Research, Syracuse University, 426 Eggers Hall, Syracuse, NY 13244-1020, USA. E-mail: whorrace@maxwell.syr.edu

(see Aigner *et al.*, 1977; Battese and Coelli, 1988). Of course, the truncated distribution of a vessel's technical inefficiency is not a point estimate of technical inefficiency per se, so the stochastic frontier literature recommends the mean of the truncated normal distribution as a point estimate for inefficiency. All the aforementioned papers use this point estimate in assessing the inefficiency. Recently, Horrace (2005) has argued that the mean of the truncated normal distribution is a poor estimate of technical inefficiency, since the mean is only one characterization of an otherwise complicated distribution. Horrace recommends calculating 'the probability that a vessel is efficient or inefficient within the sample' using a multivariate truncated normal distribution, based on the estimated mean and variance of the underlying truncated normal distribution for each vessel. These probabilities can then be used to identify vessels that are efficient or inefficient with high or low probabilities. In particular, the procedures can be used to identify a *single* vessel with a high probability of being efficient or inefficient. However, it can be shown that with reasonably large, homogeneous samples of fishing vessels a *single* efficient or inefficient vessel is not forthcoming. In this case there is no inference on vessel technical efficiency.

The purpose of this paper is to extend the probabilistic results of Horrace (2005) to construct non-empty *subsets* of minimal cardinality with high probability of being efficient or inefficient. (This is a new construction, at least outside the Bayesian part of this literature.) That is, given a sample of fishing vessels and estimates of each vessel's probability of being efficient or inefficient, the goal is to identify the smallest, non-empty subsets of vessels that will contain the most and least efficient vessels with high probability (low error rate). This is markedly different from the results of Horrace (2005), which can only identify a *single* vessel as being efficient or inefficient with high probability, and which may often result in 'no inference'. Consequently, the inference results developed here are more practical, since the 'non-empty' feature ensures that inference can always be performed. This notion of a non-empty subset is in keeping with the ranking and selection literature developed by Gupta (1965), and was first applied to the stochastic frontier model by Horrace and Schmidt (2000).

This paper also provides a practical application of the Horrace (2005) results to the North Atlantic herring fleet with an extension to the case of non-constant variance of the distributions of technical inefficiency. We demonstrate empirically that under a non-constant variance assumption the usual point estimate of technical inefficiency (based strictly on the mean of its distribution) may be an inappropriate measure with regard to any particular vessel being efficient (inefficient) with high probability. It is shown that while a large (small) point estimate of the mean of technical inefficiency implies a small (large) probability that a vessel is efficient when the variance of the underlying normal distribution is constant across vessels, this may not be the case when the variance is allowed to vary across vessels. In particular, unbalanced panels will produce erroneous inferences of relative efficiencies, because they necessarily imply heterogeneity in the variance estimates. The paper demonstrates this empirically by estimating a production function for an unbalanced panel of 39 vessels in the North Atlantic herring fleet from 2000 to 2003. We show that the rankings of the usual point estimates of technical efficiency do not correspond to Horrace's 'probability of being most efficient'.

Another empirical contribution of the paper is that the production function is estimated in two ways. Following Schnier *et al.* (2006), we estimate and perform inference on technical efficiency based on a homogeneous production function across all 39 vessels and based on a heterogeneous production function across different classes of vessels using a latent class model. Inconsistencies between our methods and the usual method of ranking the mean of technical efficiency of each vessel exist under the homogeneous production function assumption. In general, the technical

inefficiency rank statistics do not correspond to the ranks of the probabilities of vessels being most or least efficient. The inconsistency across methodologies may imply that the homogeneous production assumption is wrong (although a variety of other misspecifications may be driving this empirical result). Once the assumption is relaxed, and the production function is allowed to vary across classes of vessels, our method and the ranking of the mean efficiencies produce results that are more consistent with each other. This result underscores the importance of assessing heterogeneity in resource production functions when the mean of the inefficiency distribution is used as a point estimate for inefficiency.

The following section of the paper describes the stochastic frontier model estimated under the homogeneous and heterogeneous production function assumptions (see Schnier *et al.*, 2006), as well as the rank probability methods outlined in Horrace (2005); it also details construction of the non-empty subsets. Section 3 describes the data and discusses the heterogeneous estimation methodology of El-Gamal and Grether (1995, 2000) used in our analysis. The empirical results are outlined in Sections 4 and 5 for the homogeneous and heterogeneous production functions, respectively. The final section summarizes our findings and proposes areas for additional research.

2. PARAMETRIC FRONTIER MODELS

Battese *et al.* (1989) and Battese and Coelli (1995) formulate a stochastic production function for unbalanced panel data as follows:

$$y_{it} = f(x_{it}; \beta) \exp\{\varepsilon_{it}\} \quad (1)$$

where y_{it} is the dependent variable measured for vessel $i = 1, \dots, n$ in time period $t = 1, \dots, T_i$, x_{it} is a vector of input variables for vessel i in time period t , β is a parameter vector, and ε_{it} is the error term. The error term is specified as follows:

$$\varepsilon_{it} = v_{it} - \eta_i \quad (2)$$

where v_{it} is an independently and identically distributed $N(0, \sigma_v^2)$ random variable, and η_i is a one-sided, non-negative, vessel-specific error term distributed as the truncation below zero of a $N(\mu, \sigma_\eta^2)$ random variable, representing the technical inefficiency of vessel i . The likelihood function is (Battese *et al.*, 1989; Battese and Coelli, 1995)

$$\begin{aligned} L(\theta; y) = & -\frac{1}{2} \left(\sum_{i=1}^n T_i \right) \ln(2\pi) - \frac{1}{2} \sum_{i=1}^n (T_i - 1) \ln[(1 - \gamma)\sigma_\eta^2] \\ & - \frac{1}{2} \sum_{i=1}^n \ln\{\sigma_\eta^2[1 + (T_i - 1)\gamma]\} \\ & - n \ln[1 - \Phi(-z)] + \sum_{i=1}^n \ln[1 - \Phi(-z_i^*)] - \frac{1}{2} n z^2 \\ & - \frac{1}{2} (y_i - x_i \beta)' (y_i - x_i \beta) [(1 - \gamma)\sigma_\eta^2] + \frac{1}{2} \sum_{i=1}^n z_i^{*2} \end{aligned}$$

where

$$z = \frac{\mu}{(\sigma_S^2 \gamma)^{0.5}}, \quad z_i^* = \frac{\mu(1 - \gamma) - T_i \gamma (\bar{y}_i - \bar{x}_i \beta)}{(\gamma(1 - \gamma) \sigma_S^2 [1 + (T_i - 1) \gamma])^{0.5}} \quad (3)$$

The parameters for estimation are β , $\gamma = \sigma_V^2 / \sigma_S^2$, μ and $\sigma_S^2 = (\sigma_U^2 + \sigma_V^2)$.¹

The conditional mean of the vessel-specific error term, η_i , can be used to determine the level of technical efficiency possessed by each vessel. The traditional technical efficiency of each vessel is (Battese *et al.*, 1989; Battese and Coelli, 1992, 1993):

$$TE_i = E[\exp\{-\eta_i | \varepsilon_{it}\}] = \frac{1 - \Phi \left[\sigma_i^* - \left(\frac{\mu_i^*}{\sigma_i^*} \right) \right]}{1 - \Phi \left(\frac{-\mu_i^*}{\sigma_i^*} \right)} \exp \left\{ -\mu_i^* + \frac{1}{2} \sigma_i^{*2} \right\} \quad (4)$$

where

$$\mu_i^* = \frac{\mu \sigma_V^2 - \sigma_U^2 T_i \bar{E}_i}{\sigma_V^2 + T_i \sigma_U^2} \quad (5)$$

$$\bar{E}_i = \frac{1}{T_i} \sum_{t=1}^{T_i} \hat{\varepsilon}_{it} \quad (6)$$

$$\sigma_i^{*2} = \frac{\sigma_U^2 \sigma_V^2}{\sigma_V^2 + T_i \sigma_U^2} \quad (7)$$

The $\hat{\varepsilon}_{it}$ is the residual from the maximum likelihood estimate. Here (5) and (7) are the mean and variance of a normal distribution, which (when truncated at zero) represent the conditional distribution of technical inefficiency of vessel i . That is, the truncated normal distribution of η_i conditional on ε_i is given by the truncation at zero of an $N(\mu_i^*, \sigma_i^{*2})$. Equation (4) is the mean of a monotonic transformation of this truncated distribution. It has been argued in Horrace (2005) that (4) is a misleading measure of technical efficiency and that study recommends the empirical probabilities

$$F_k(\mu_i^*, \sigma_i^{*2}) = F_k = \Pr\{\eta_i \leq \eta_k, \forall i \neq k\} \quad (8)$$

and

$$F_k^*(\mu_i^*, \sigma_i^{*2}) = F_k^* = \Pr\{\eta_i \geq \eta_k, \forall i \neq k\} \quad (9)$$

(see Horrace, 2005, equations 4 and 5, for explicit formulae for these probabilities based on the normal-half-normal specification). The probabilities F_k and F_k^* are readily calculable based on the estimates in (5) and (7), and correspond to ‘the probability that vessel k is most inefficient (least efficient)’ and ‘the probability that vessel k is least inefficient (most efficient)’, respectively. In particular, if F_k^* is relatively large, then one can say that vessel k is least inefficient (most efficient) with high probability. A shortcoming of this concept is that for large values of n and

¹ In the case of heterogeneous production each class within the latent class regressions possesses a separate parameter vector, β_h .

vessels of roughly the same efficiency this probability, F_k^* , will necessarily be small. That is, it will typically be difficult for any one vessel to be least efficient with high probability. To see this, one need only recognize that the events in (8) and (9) are partitions of the probability space (see Horrace, 2005, p. 341). That is, the events do not overlap, and the sum of the F_k over k is equal to 1. For instance, if there are three vessels (of many) that are technically efficient (on the efficient frontier), then the probability that any one of these three vessels is least inefficient is necessarily less than 1/3, a fairly low probability indeed.

Consider an alternative inferential approach developed here. The goal is not to identify a *single* vessel as least inefficient or most inefficient with high probability, but to identify a non-empty subset of vessels of minimum cardinality that contains the least and most inefficient vessel with high probability. This style of inference is more typical of the ranking and selection literature (e.g., Gupta, 1965) than is the inference of Horrace (2005). Let the ranked probabilities be

$$F_{(n)} > F_{(n-1)} > \dots > F_{(1)} \text{ and } F_{\langle n \rangle}^* > F_{\langle n-1 \rangle}^* > \dots > F_{\langle 1 \rangle}^*$$

Notice the notational subtlety that (k) is distinct from $\langle k \rangle$ in the rankings. Due to the multivariate nature of the probabilities, it should be noted that $F_k \neq 1 - F_k^*$, in general. Consequently, (1) does not necessarily equal (n) , and $\langle 1 \rangle$ does not necessarily equal $\langle n \rangle$. If there are enough significant digits in the calculation of μ_i^* , σ_i^{*2} , F_k , and F_k^* , ties in the rankings should not be a factor. However, any tie-breaking rule (including randomization) will not compromise the inference that follows. It should also be noted that these ranked probabilities may not correspond to the ranked technical efficiencies of equation (4). This feature of the calculations is made clear in the empirical analyses that follow, and it underscores the importance of not confusing the probabilities in equations (8) and (9) with technical efficiency scores. Define the subsets:

$$\zeta_{\max} = \{(n), (n - 1), \dots, (k)\}, \text{ where } (k) \text{ is the largest value of } (j) \text{ such that } \sum_{i=(j)}^{(n)} F_i \geq 1 - \lambda$$

and

$$\zeta_{\min} = \{\langle n \rangle, \langle n - 1 \rangle, \dots, \langle k \rangle\}, \text{ where } \langle k \rangle \text{ is the largest value of } \langle j \rangle \text{ such that } \sum_{i=\langle j \rangle}^{\langle n \rangle} F_i^* \geq 1 - \alpha$$

and where $\lambda \in (0, 0.5)$ and $\alpha \in (0, 0.5)$ are error rates that are important in the sequel. The subset $\zeta_{\max} \subset \{1, \dots, n\}$ contains the indices associated with the largest probabilities F_j , while the subset $\zeta_{\min} \subset \{1, \dots, n\}$ contains the indices associated with the largest probabilities F_j^* . Since the subsets ζ_{\max} and ζ_{\min} contain at least the index (n) and $\langle n \rangle$, respectively, inference on the minimal and maximal η_i (the least and most efficient vessel) is assured (ζ_{\max} and ζ_{\min} are non-empty).

Let Ψ_{\max} be the selection rule associated with the set ζ_{\max} , and let Ψ_{\min} be the selection rule associated with the set ζ_{\min} . Let the ranked inefficiency terms be $\eta_{[n]} > \eta_{[n-1]} > \dots > \eta_{[1]}$, so that $[n]$ corresponds to the index of the most inefficient vessel in the sample and $[1]$ corresponds to the index of the least inefficient vessel in the sample. Further, let correct selections CS and CS^* be defined as events: $CS = \{[n] \in \zeta_{\max}\}$ given Ψ_{\max} and $CS^* = \{[1] \in \zeta_{\min}\}$ given Ψ_{\min} , respectively. That is, a correct selection occurs when ζ_{\max} and ζ_{\min} contain the indices associated with the most $[n]$ and least $[1]$ inefficient vessels, respectively. Then the following result is true.

Result 1 Given the selection rules Ψ_{\max} and Ψ_{\min} , ζ_{\max} and ζ_{\min} are non-empty subsets of minimal cardinality such that

$$\Pr\{CS\} = \Pr\{[n] \in \zeta_{\max}\} \geq 1 - \lambda \text{ and } \Pr\{CS^*\} = \Pr\{[1] \in \zeta_{\min}\} \geq 1 - \alpha \quad \square$$

The result follows directly from the way that the subsets are defined, but a formal proof is provided in the Appendix. The result implies that ζ_{\max} and ζ_{\min} contain the indices associated with the most and least inefficient vessels, respectively, with pre-specified confidence levels $1 - \lambda$ and $1 - \alpha$, respectively.² The selection rules Ψ_{\max} and Ψ_{\min} nest the selection rules in Horrace (2005). That is, the Horrace (2005) subsets are contained in ζ_{\max} and ζ_{\min} . The inference hinges on independence of the technical inefficiency distributions. In principle, this independence assumption can be relaxed; however, calculation of the probabilities F_k and F_k^* becomes cumbersome even for relatively small values of n . It should also be noted that these results do not require homogeneity of the production function implied by equation (1). In fact, any specification of a conditional mean production function (homogeneous or not) that produces estimates of μ_i^* and σ_i^* is acceptable for use with Result 1.

Notice that the *most inefficient* vessel in the sample is equivalent to the *least efficient* vessel in the sample, so it can be said that ζ_{\max} contains the *least efficient* vessel with high probability. It may seem counter-intuitive to associate the ‘max’ subscript of ζ_{\max} with the ‘minimally efficient vessel’, but one only need remember that the minimal efficiency is equivalent to *maximal* inefficiency. In what follows we attempt to be consistent with the notation and couch all discussion on ζ_{\max} in terms of maximal *inefficiency* and all discussion on ζ_{\min} in terms of minimal *inefficiency*.

We should also mention that there is a Bayesian inference literature that has grown out of the stochastic frontier literature as well. These techniques either directly or indirectly provide inference on relative ranks using Bayesian sampling techniques and are a viable alternative to the results presented here; for example, see Atkinson and Dorfman (2005), Fernandez *et al.* (2002), Tsionas (2002), Kim and Schmidt (2000), and Koop *et al.* (1997).

3. DATA AND METHODOLOGY

We exploit the selection rules in a new application to the North Atlantic herring fleet. To estimate the stochastic frontier model we use logbook data for the Northeast Atlantic herring fleet provided by the National Marine Fisheries Service (NMFS). Logbook data contain information on where a vessel fishes, what gear they employ, the time and date of departure and return, their home port, the number of crew members on board, and a list of vessel characteristics (length, gross registered tonnage, horsepower, and hold capacity). The data set is identical to that used by Schnier *et al.* (2006) and it contains 2894 observations for 39 vessels over the years 2000–2003. The data are reported at two levels of spatial resolution: the centroid of the latitude and longitude coordinate of the statistical reporting area a vessel fishes and the macro-region within which this statistical reporting is contained. The first level of spatial resolution is used to determine the distance that vessels travel to fish to calculate the hours that they fished. The second level of spatial resolution is

² It should be noted that the subsets are not necessarily disjoint. A referee pointed out that it would be useful to modify the selection rules so that the subsets are guaranteed to be disjoint; this is left for future research.

used to control for the unobserved stock density within the area fished and the seasonal migration patterns of herring. In particular, we create dummy variables to indicate whether vessels fished inshore or offshore during particular seasons. Each observation in the logbook data set represents a single trip and is the most accurate data available on the Atlantic herring fleet.

The econometric specifications used to illustrate our inference methods are identical to those in Schnier *et al.* (2006). The dependent variable used within the regressions is the total catch, C_{it} , in tonnes of fish for each of the logbook data entries. The logbook data contain the latitude and longitude coordinates of the vessel's home port as well as the statistical reporting area fished. Using these coordinates Schnier *et al.* were able to estimate the distance traveled by each vessel. In addition, the temporal and spatial resolution of the logbook data allows them to determine the distance traveled and, hence, the hours each vessel spent steaming to the fishing locations (assuming an average vessel speed of 12 knots).³ They were then able to recover an estimate of total hours spent fishing (gear deployed) on each trip for each vessel. They estimate the following heterogeneous production function:

$$\begin{aligned} \ln(C_{it}) = & \beta_{0|h} + \beta_{1|h} * \ln(\text{GRT}_i) + \beta_{2|h} * \ln(\text{HP}_i) + \beta_{3|h} * \ln(\text{Crew}_{it}) + \beta_{4|h} * \ln(\text{Hours}_{it}) \\ & + \beta_{5|h} * \ln(\text{GRT}_i) * \ln(\text{Crew}_{it}) + \beta_{6|h} * \ln(\text{Crew}_{it}) * \ln(\text{Hours}_{it}) + \beta_{7|h} * \text{DumNoCrew}_{it} \\ & + \beta_{8|h} * \text{DumSpWntInshore}_{it} + \beta_{9|h} * \text{DumSpWntOffshore}_{it} \\ & + \beta_{10|h} * \text{DumSumFallOffshore}_{it} + v_{it} - \eta_i \end{aligned} \quad (10)$$

where $h = 1, \dots, H$ indexes production tiers, GRT is the vessel's gross registered tonnage, HP is the vessel's horsepower, Crew_{it} is the number of crew members employed, Hours_{it} is the total hours that the gear was deployed during the trip, and DumNoCrew_{it} is a dummy variable indicating whether or not the number of crew members on board the vessel i were observed in time period t . In the case that no crew members were observed, we substitute the mean number of crew members in the data set for the missing value. The variables $\text{DumSpWntInshore}_{it}$, $\text{DumSpWntOffshore}_{it}$, and $\text{DumSumFallOffshore}_{it}$ are dummy variables indicating whether vessel i fished inshore during the spring or winter, offshore during the spring or winter or offshore during the summer or fall in time period t , respectively. The dummy variables control for the respective inshore and offshore seasons and for the stock abundances present during these time periods.⁴ Equation (10) is estimated assuming both homogeneous and heterogeneous production technologies. In the homogeneous case, this means that all vessels employ the same short-run production functions (e.g., $\beta_{4|h} = \beta_4$ for all h). Under heterogeneous production, each tier, h , within the fleet possesses a tier-specific production function captured by the variability of the coefficients over h .

To estimate the heterogeneous production function the El-Gamal and Grether estimation classification (EC) algorithm is used (El-Gamal and Grether, 1995, 2000). In addition to Schnier *et al.*'s use of the EC algorithm, this method has recently been used by El-Gamal and Inanoglu (2005) to investigate heterogeneity in the Turkish banking sector. The EC algorithm groups vessels into a pre-specified number of tiers, H , and estimates the parameter vector, $\theta_h = [\beta_{0|h}, \dots, \beta_{11|h}, \gamma, \mu, \sigma_s^2]$,

³ Average vessel speed was provided by contacts at Woods Hole Oceanographic Institute. For a more detailed description of the data utilized and the methodology used to obtain the hours fished see Schnier *et al.* (2006).

⁴ It is acknowledged that the decision of 'where to fish' is part of 'skipper skill'. Since we are controlling for this in the regression (inshore or offshore), the remaining technical inefficiency will not include this location component. However, inshore and offshore locations are very large areas, so the inefficiency will still capture the important decision of where to fish *within* these inshore or offshore areas.

for each tier assuming γ , μ , and σ_s^2 are held constant across all tiers.⁵ To determine which vessels belong to each of the H tiers, each vessel's contribution to the likelihood function is the maximum of the joint likelihood of all their observations, given $\Theta = [\theta_1, \dots, \theta_H]$.⁶ The econometric results for the homogeneous production technology are presented in Table I. The heterogeneous production results are also in Table I.⁷ Schnier *et al.* (2006) proposed the use of the heterogeneous production functions to investigate whether or not 'highliners' can be better represented by productivity measures and to illustrate that policy recommendations regarding excess capacity may be erroneous if they are based on the traditional homogeneous production function assumption. For a complete discussion of the results and the implied elasticities (given the interactions in the model), see Schnier *et al.* (2006).

4. HOMOGENEOUS PRODUCTION FUNCTION RESULTS

The efficiency estimation results, assuming a homogeneous production technology, are contained in Table II for each vessel. The first column of the table contains the vessel number, i . The second column contains the variance of the technical efficiency distribution of vessel i before truncation at zero (equation (7)). The third column contains the mean of the technical efficiency distribution for vessel i before truncation at zero (equation (5)). The fourth column contains the technical efficiency estimate for vessel i (equation (4)) based on the mean of the efficiency distribution after truncation. Notice that the results imply that vessels 3 and 33 are most efficient (TE = 0.956 and 0.933, respectively), and vessels 14 and 38 are least efficient (TE = 0.016 and 0.007, respectively). This is technical efficiency in an *absolute* sense, relative to some out-of-sample, absolute standard. These are the parametric stochastic frontier efficiency results that one typically sees. Notice that for the least efficient vessels the technical efficiencies are surprisingly small (this improves with the heterogeneous estimation, so this is, in part, why we prefer the heterogeneous model). Columns 5 and 6 of the table are the new results from Horrace (2005), with F_i equal to the probability that vessel i is most inefficient and F_i^* equal to the probability that vessel i is least inefficient.⁸ These efficiency results are *relative* to other vessels in the sample (relative inefficiency as opposed to absolute inefficiency).⁹ In this (and all the tables that follow) the vessels are ranked on F_i^* .

⁵ Holding γ , μ and σ_s^2 constant across the tiers is a necessary assumption because the econometric model is 'ill-posed' if these parameters vary across the tiers (El-Gamal and Inanoglu, 2005).

⁶ The joint likelihood is expressed as (Schnier *et al.*, 2006)

$$\ln[L(y_{it}; x_{it} | \Theta; H)] = \sum_{i=1}^N \arg \max_h \sum_{t=1}^{T_i} \ln(L(y_{it}; x_{it} | \theta_h))$$

where $L(\cdot)$ is the likelihood function for the stochastic frontier model outlined in equation (3). This likelihood function may possess many local maxima. Therefore, 500 random starting points were used to obtain the maximum value using the constrained maximum likelihood (CML) algorithm in GAUSS for each of the iterations.

⁷ Given that the number of segments, H , is predetermined, likelihood ratio tests, Bayesian information criterion (BIC) and Akaike information criterion (AIC) were used to determine the optimal number of segments to use within the estimation. For a more detailed description of the estimation procedure see Schnier *et al.* (2006).

⁸ These probabilities are calculated with the 'rectangle' algorithm in Horrace (2005, footnote 6, p. 349).

⁹ It is, indeed, the relative nature of these efficiency measures that provide the extra information to identify variance differences across vessels (heterogeneity) in the sequel.

Table I. Stochastic frontier results (Schnier *et al.*, 2006)^a

Variable	Homogenous	Tier 1	Tier 2	Tier 3
Constant	-2.2913 (-1.62)	-16.0335** (-11.03)	-10.1388** (-3.88)	1.8865* (1.67)
GRT	0.2653* (1.71)	-0.5028** (-3.26)	2.7700** (6.91)	0.1658 (1.42)
HP	0.7781** (3.57)	2.8326** (10.85)	0.1254 (0.38)	0.3249** (2.16)
Crew	0.2679 (0.62)	0.4373 (1.06)	9.8811** (4.06)	-0.4213 (-1.05)
Hours	0.1130 (1.20)	0.4405** (2.99)	-0.2826** (-2.28)	-0.3380** (-2.24)
GRT * Crew	0.0184 (0.17)	-0.0146 (-0.13)	-2.0462** (-4.36)	0.0845 (1.10)
Crew * Hours	-0.0668 (-1.04)	-0.1457 (-1.32)	0.2823** (3.08)	0.1414 (1.49)
DumNoCrew	-0.1314** (-2.95)	0.0081 (0.08)	-0.1752** (-2.86)	-0.0605 (-0.79)
DumSpWntInshore	-0.3221** (-6.93)	-0.0486 (-0.53)	-0.0183 (-0.22)	-0.5759** (-9.25)
DumSpWntOffshore	0.3829** (7.70)	0.6814** (7.24)	0.3106** (3.94)	-0.1518* (-1.66)
DumFallSumOffshore	0.2094** (4.51)	-0.0314 (-0.31)	0.3870** (6.88)	-0.2274* (-1.83)
γ	0.9597** (25.19)	—	—	0.8807** (13.88)
σ_S^2	10.3535 (1.06)	—	—	3.2835* (1.88)
μ	-3.9023 (-0.61)	—	—	-6.9310 (-1.34)
Number of vessels	39	13	12	14
Mean log-likelihood	-1.00957	—	—	-0.96396

^a For a more detailed discussion of the implied elasticities and marginal products see Schnier *et al.* (2006).

GRT, gross registered tonnage; HP, horsepower.

** Indicates significant at the 95% level; * indicates significant at the 95% level; *t*-statistics in parentheses.

The results in Table II are compelling. Even though vessels 30 and 35 possess relatively high values of ranked TE (0.864 and 0.809, respectively), they are least inefficient with very low probability, based on F_i^* . Also, notice that vessel 8 possesses a relatively low TE score (0.648) but is efficient with more than three times the probability of vessels 30 and 35 (compare 0.0304 to 0.0029 and 0.0104). Why this contradictory result? The answer lies in the distribution of technical inefficiency for the vessels in question. Notice that prior to truncation the distributions of technical inefficiency for vessels 30 and 35 possesses relatively low variance ($\sigma_{30}^{*2} = 0.001$ and $\sigma_{35}^{*2} = 0.07$, respectively). Due to the low variances, we can reject the hypothesis that these two vessels are efficient, relative to vessel 8, which has a high variance, $\sigma_8^{*2} = 0.204$. For these results, the variance of the distribution after truncation, the variance of interest in these comparisons, is an increasing function of the variance before truncation.¹⁰ Therefore, after truncation the distributions of vessels

¹⁰ For the formula for the variance after truncation see Horrace (2005, equation 2). It is not clear whether this a general theoretical result or an artifact of the data. Investigation of this result is left for future research.

Table II. Homogeneous vessel efficiency results sorted on F_i^*

Vessel i	σ_i^{*2}	μ_i^*	TE_i	F_i	F_i^*
3	0.001	0.032	0.956	0	0.5086
33	0.002	0.065	0.933	0	0.2086
12	0.137	-0.142	0.794	0	0.0856
34	0.024	0.210	0.795	0	0.0406
8	0.204	0.278	0.648	0	0.0304
19	0.042	0.272	0.746	0	0.0302
6	0.103	0.316	0.684	0	0.0288
10	0.137	0.335	0.660	0	0.0272
27	0.400	0.273	0.586	0	0.0257
35	0.007	0.215	0.809	0	0.0104
30	0.001	0.147	0.864	0	0.0029
7	0.400	1.508	0.262	0	0.0012
18	0.059	1.128	0.334	0	0
20	0.052	1.264	0.290	0	0
23	0.137	2.103	0.131	0	0
21	0.012	0.674	0.513	0	0
31	0.001	0.301	0.741	0	0
1	0.204	2.863	0.063	0	0
13	0.009	0.680	0.509	0	0
26	0.007	0.599	0.551	0	0
36	0.010	0.773	0.464	0	0
29	0.007	0.670	0.514	0	0
16	0.008	0.761	0.469	0	0
17	0.042	2.135	0.121	0	0
32	0.103	3.849	0.022	0.0022	0
5	0.035	2.623	0.074	0	0
24	0.009	1.549	0.213	0	0
14	0.069	4.188	0.016	0.0129	0
25	0.007	1.395	0.249	0	0
28	0.028	2.975	0.052	0	0
22	0.006	1.555	0.212	0	0
37	0.042	3.887	0.021	0.0001	0
4	0.022	2.907	0.055	0	0
15	0.001	0.965	0.381	0	0
39	0.014	2.598	0.075	0	0
38	0.042	4.928	0.007	0.9848	0
2	0.001	1.423	0.241	0	0
9	0.003	1.845	0.158	0	0
11	0.012	3.566	0.028	0	0

TE_i is the technical efficiency of vessel i .

F_i is the probability that vessel i is least efficient.

F_i^* is the probability that vessel i is most efficient.

$\zeta_{\max} = \{38\}$ at 95% level.

$\zeta_{\min} = \{3, 6, 8, 10, 12, 19, 33, 34\}$ at 95% level.

30 and 35 also have low variance. This distinction underscores the fallacy of using the mean of the distribution of technical efficiency as an estimate of technical efficiency in the sample. The multivariate probability F_i^* is simply a more appropriate measure. It also underscores the importance of accounting for heterogeneity (different σ_i^{*2}) when this is a possibility.¹¹ This is echoed in the

¹¹ Heterogeneity is generally not identified in cross-section parametric stochastic frontier models, but only in panel data models.

inconsistent rankings of vessels 8 and 19 and (to a lesser extent) vessels 12 and 34. In all cases, variance differences of the vessels are driving the ranking differences between the TE score and F_i^* .

Using Result 1, we conclude that at the 95% level: $\zeta_{\max} = \{38\}$ and $\zeta_{\min} = \{3, 6, 8, 10, 12, 19, 33, 34\}$. Therefore with probability 0.95 the most inefficient vessel is 38, and the least inefficient vessel is either vessel 3, 6, 8, 10, 12, 19, 33, or 34. Notice, again, that vessel 30 is not even in the subset of the best (least inefficiency), while vessel 8 is. However, had we focused attention only on the TE_i , we would have come to the fallacious conclusion that vessel 30 is relatively efficient, while in a probabilistic sense it is not likely to be, when its distribution of technical efficiency is compared to those of other vessels in the sample (i.e., when F_i^* is used as a means of comparison). A valid question is whether or not the fallacious conclusion is due to the overly simplistic estimate, TE_i , or to the fact that a homogeneous production function has been assumed for all 39 vessels. Of course, for this empirical exercise we cannot answer this definitively, but the empirical results that follow suggest that it may be due to the homogeneous production specification.

Using likelihood ratio tests, Bayesian information criterion (BIC) and Akaike information criterion (AIC) in the context of the EC algorithm, Schnier *et al.* (2006) conclude that three different (heterogeneous) production functions fit the data better than a single homogeneous function. Hence, we divide the sample of vessels into the three segments of Schnier *et al.*, which correspond to the three different production functions based on the latent class regression model and estimated using the EC algorithm. Then, using the results on μ_i^* and σ_i^* in Table II (based on a homogeneous production function), we recalculate relative probabilities F_i and F_i^* based on the heterogeneous tiers. As such, the technical efficiency results are relative to the sample segments proposed by Schnier *et al.* These are contained in Tables III–V for each subset. These tables are read identically to Table II, and the only values that have changed are the relative probabilities, F_i and F_i^* . The importance of the variance is underscored in these tables as well. For example, in Table V, the relatively high variability of vessel 34 does not preclude it from being one of the least inefficient vessels based on the F_i^* , while (again) the low variability of vessels 30 and 35 do. The usual TE_i estimates do not capture these heterogeneity effects, and one would have erroneously concluded that vessel 30 is efficient relative to 34. Finally in all three tables, there are still very few efficient vessels, and the inefficient vessels have extremely low efficiency scores (e.g., $TE_i = 0.007$ for vessel 38 in Table III). One could view these low scores as counter-intuitive, but this was merely a thought experiment to see if forcing the homogeneous results into the tiers implied by the heterogeneous results improves inference over the purely homogeneous case. In general, it does not. However, things are improved when we move to the heterogeneous production function, the empirical focus of the next exercise.

5. HETEROGENEOUS PRODUCTION FUNCTION RESULTS

The heterogeneous production results of Schnier *et al.* (2006) are reproduced in the last three columns of Table I. Again, the EC algorithm divides the sample of vessels into three segments, which correspond to three different production functions, based on the latent class regression. We reproduce the latent class inefficiency results of Schnier *et al.* for each segment in Tables VI–VIII. In Table VI, vessels 1, 2, 5, 10, 11, 14, 27, 39, and 37 are the relatively most efficient (least inefficient) with probability at least 0.95. Now many vessels are operating on the frontier and the least efficient vessel (vessel 38) is now 29.8% efficient (versus 0.7% efficient in the homogeneous

Table III. Homogeneous vessel efficiency results, sorted on F_i^*

Vessel i	σ_i^{*2}	μ_i^*	TE $_i$	F_i	F_i^*
10	0.137	0.335	0.660	0	0.5987
27	0.400	0.273	0.586	0	0.4012
2	0.001	1.423	0.241	0	0.0001
1	0.204	2.863	0.063	0	0
17	0.042	2.135	0.121	0	0
5	0.035	2.623	0.074	0	0
28	0.028	2.975	0.052	0	0
39	0.014	2.598	0.075	0	0
4	0.022	2.907	0.055	0	0
14	0.069	4.188	0.016	0.0130	0
37	0.042	3.887	0.021	0.0001	0
11	0.012	3.566	0.028	0	0
38	0.042	4.928	0.007	0.9868	0

F_i is the probability that vessel i is least efficient.

F_i^* is the probability that vessel i is most efficient.

$\zeta_{\max} = \{38\}$ at 95% level.

$\zeta_{\min} = \{10, 27\}$ at 95% level.

Table IV. Homogeneous vessel efficiency results, sorted on F_i^*

Vessel i	σ_i^{*2}	μ_i^*	TE $_i$	F_i	F_i^*
12	0.137	-0.142	0.794	0	0.4283
19	0.042	0.272	0.746	0	0.2334
6	0.103	0.316	0.684	0	0.1713
8	0.204	0.278	0.648	0	0.1590
7	0.400	1.508	0.262	0.0005	0.0071
21	0.012	0.674	0.513	0	0.0006
13	0.009	0.680	0.509	0	0.0003
18	0.059	1.128	0.334	0	0.0001
15	0.001	0.965	0.381	0	0
24	0.009	1.549	0.213	0	0
32	0.103	3.849	0.022	0.9995	0
9	0.003	1.845	0.158	0	0

F_i is the probability that vessel i is least efficient.

F_i^* is the probability that vessel i is most efficient.

$\zeta_{\max} = \{32\}$ at 95% level.

$\zeta_{\min} = \{6, 8, 12, 19\}$ at 95% level.

results). These results underscore the importance of accounting for heterogeneity in the production functions of fishing vessels. Oddly, vessel 14 is now in contention for high efficiency (least inefficiency) in Table VI, even though it was close to the bottom in terms of its probability of being most inefficient in the homogeneous case (Tables II and III). The main difference is that with the assumption of a homogeneous production function in Tables II and III, the mean of the distribution of technical inefficiency for vessel 14 (prior to truncation) is large (4.188), while under a heterogeneous production function assumption it is small (-0.077). This makes sense, since differences in the specification of the conditional mean production function are more likely to manifest themselves as differences in the conditional mean of technical efficiency than as

Table V. Homogeneous vessel efficiency results, sorted on F_i^*

Vessel i	σ_i^{*2}	μ_i^*	TE $_i$	F_i	F_i^*
3	0.001	0.032	0.956	0	0.6545
33	0.002	0.065	0.933	0	0.2767
34	0.024	0.210	0.795	0	0.0501
35	0.007	0.215	0.809	0	0.0138
30	0.001	0.147	0.864	0	0.0050
20	0.052	1.264	0.290	0.0124	0
23	0.137	2.103	0.131	0.9192	0
31	0.001	0.301	0.741	0	0
26	0.007	0.599	0.551	0	0
36	0.010	0.773	0.464	0	0
29	0.007	0.670	0.514	0	0
16	0.008	0.761	0.469	0	0
25	0.007	1.395	0.249	0.0037	0
22	0.006	1.555	0.212	0.0646	0

F_i is the probability that vessel i is least efficient.

F_i^* is the probability that vessel i is most efficient.

$\zeta_{\max} = \{22, 23\}$ at 95% level.

$\zeta_{\min} = \{3, 33, 34\}$ at 95% level.

Table VI. Heterogeneous vessel efficiency results, sorted on F_i^*

Vessel i	σ_i^{*2}	μ_i^*	TE $_i$	F_i	F_i^*
39	0.013	0.030	0.905	0	0.1665
27	0.345	-1.553	0.844	0.0008	0.1280
2	0.001	0.074	0.928	0	0.1268
10	0.125	-0.360	0.841	0.0003	0.1132
14	0.064	-0.077	0.846	0	0.1058
5	0.032	0.040	0.859	0	0.1028
11	0.011	0.094	0.881	0	0.0998
37	0.039	0.126	0.817	0	0.0633
1	0.183	0.034	0.725	0.0104	0.0458
4	0.020	0.242	0.781	0	0.0233
17	0.039	0.303	0.731	0.0004	0.0200
28	0.026	0.387	0.685	0.0004	0.0049
38	0.039	1.230	0.298	0.9876	0

F_i is the probability that vessel i is least efficient.

F_i^* is the probability that vessel i is most efficient.

$\zeta_{\max} = \{38\}$ at 95% level.

$\zeta_{\min} = \{1, 2, 5, 10, 11, 14, 27, 39, 37\}$ at 95% level.

differences in the conditional variance. Notice that, when the production function (conditional mean) changes, the conditional variance of vessel 14 does not change all that much: 0.069 (in Tables II and III) versus 0.064 (in Table VI).

Generally speaking, the TE $_i$ results in Tables VI–VIII correspond to the inferences implied by the ζ_{\min} and ζ_{\max} results (with a few exceptions). In particular, the rank ordering at the top and bottom of the tables correspond, whether we focus on TE $_i$ or F_i^* . For instance, in Table VI and VII, the subsets of least inefficiency, $\zeta_{\min} = \{1, 2, 5, 10, 11, 14, 27, 39, 37\}$ and

Table VII. Heterogeneous vessel efficiency results, sorted on F_i^*

Vessel i	σ_i^{*2}	μ_i^*	TE_i	F_i	F_i^*
19	0.039	-0.182	0.903	0	0.2347
21	0.011	0.041	0.907	0	0.2026
12	0.125	-0.518	0.863	0.0001	0.1635
7	0.345	-1.129	0.814	0.0040	0.1168
13	0.009	0.135	0.864	0	0.0758
32	0.095	0.095	0.768	0.0018	0.0630
8	0.183	-0.047	0.743	0.0120	0.0622
6	0.095	0.126	0.758	0.0023	0.0575
18	0.055	0.332	0.705	0.0015	0.0238
24	0.009	0.967	0.382	0.2643	0
15	0.001	0.692	0.501	0	0
9	0.002	1.032	0.357	0.7140	0

F_i is the probability that vessel i is least efficient.

F_i^* is the probability that vessel i is most efficient.

$\zeta_{\max} = \{9, 24\}$ at 95% level.

$\zeta_{\min} = \{6, 7, 8, 12, 13, 19, 21, 32\}$ at 95% level.

Table VIII. Heterogeneous vessel efficiency results, sorted on F_i^*

Vessel i /Type	σ_i^{*2}	μ_i^*	TE_i	F_i	F_i^*
3/P	0.001	0.001	0.971	0	0.6047
16/T	0.008	0.011	0.930	0	0.1932
33/P	0.002	0.065	0.934	0	0.1177
34/P	0.023	0.173	0.817	0	0.0369
20/T	0.048	0.179	0.783	0	0.0364
23/T	0.125	0.604	0.555	0.0437	0.0060
35/P	0.006	0.215	0.808	0	0.0051
30/P	0.001	0.193	0.825	0	0
29/P	0.007	0.504	0.606	0	0
26/P	0.006	0.602	0.550	0	0
36/P	0.010	0.913	0.403	0.0055	0
22/T	0.006	0.836	0.435	0.0001	0
31/P	0.001	0.494	0.611	0	0
25/T	0.006	1.230	0.293	0.9506	0

TE_i reported in Schnier *et al.* (2006).

F_i is the probability that vessel i is least efficient.

F_i^* is the probability that vessel i is most efficient.

$\zeta_{\max} = \{25\}$ at 95% level.

$\zeta_{\min} = \{3, 16, 33, 34\}$ at 95% level.

$\zeta_{\min} = \{6, 7, 8, 12, 13, 19, 21, 32\}$ (respectively), correspond to vessels with high values of technical efficiency, TE_i , in proper rank order (for the most part). Results are admittedly more consistent in the heterogeneous results of Tables VI–VIII than in the homogeneous results of Table II–V, but the erroneous ‘rank conclusions’ of the technical efficiency estimates of vessels 30 and 35 in Table VIII still cannot be ignored.

According to Schnier *et al.* (2006), the vessels in tiers one (Table VI) and two (Table VII) use similar fishing technologies: mid-water trawling, where a large net is dragged behind the vessel to catch fish. Tier three (Table VIII), however, possesses a mix of two fishing techniques. Some of the vessels are mid-water trawlers, but most of them use purse seine gear, which is not dragged,

but used to encircle pods of fish. Not surprisingly, the heterogeneous efficiency results of this tier are different from those of tiers one and two, where the fishing techniques are similar across vessels. The first column of Table VIII contains the vessel number and the type of net employed: P = purse-seine gear and T = trawler. It seems that trawlers tend to be *probabilistically* most and least inefficient, when compared to the purse seine gear vessels, as they tend to be grouped at the top and bottom of the rank statistic in the table.

6. CONCLUSIONS

This paper extends the probability results of Horrace (2005) to produce non-empty subsets of minimal cardinality that contain the least and most inefficient vessels with predetermined probability. The new results are useful, because the results of Horrace (2005) may not always exist, while our inference will always exist. The application to fishing vessels was particularly relevant because it represents a case where the results imply that heterogeneous production functions are preferred to a homogeneous production function. Under homogeneity the inference implied by the usual TE_i estimate, and that implied by the non-empty subsets, were not consistent. Under heterogeneity of production, correspondence of the results was markedly improved. That is, the vessels at the top and bottom of the ranking on the probability F_i^* tended to correspond to vessels at the extreme end of the ranking on the TE_i score.

One referee pointed out that an interesting extension would be to rank the vessels on F_i^* and F_i , trimming the vessels in the subsets ζ_{\min} and ζ_{\max} from the sample. The inference could then be re-performed on the remaining subset of vessels, producing second-best (least inefficient) and second-worst (most inefficient) subsets. The trimming procedure could be repeated until the original sample was exhausted. A similar experiment was performed in Horrace *et al.* (2007) in the ranking of income inequality and poverty measures across countries. But as pointed out in that paper, the procedure does not correctly control for the overall error rate of the problem. Figuring out how one might adequately control for this error rate might prove useful in subsequent research.

Finally, the results extend beyond the model illustrated within this research. Global fisheries management has been plagued by the concern of excess capacity and overcapitalization as a result of the 'race to fish'. One policy solution to this phenomenon has been the support of vessel buybacks when fisheries policy is restructured. These buybacks are often targeted at removing the most inefficient vessels within a fleet to reduce the degree of excess capacity (Guyader *et al.*, 2004). The methodologies illustrated in this paper will allow policy makers to more concretely estimate which vessels are the most inefficient within a fleet and facilitate the development of more efficient fisheries policy, which in turn should help alleviate the pressures on the resource.

APPENDIX: PROOF OF RESULT 1

Suppose that $\zeta_{\max} = \{(n), (n-1), \dots, (j+1), (j)\}$ associated with the largest probabilities:

$$F_{(n)} > F_{(n-1)} > \dots > F_{(j+1)} > F_{(j)}$$

In particular, it is supposed that the index $(j-1)$ associated with $F_{(j-1)} < F_{(j)}$ is not an element of ζ_{\max} . Then by the continuity of η and by definition of ζ_{\max} , it is true that

$$F_{(n)} > F_{(n-1)} > \dots > F_{(j+1)} < 1 - \lambda$$

Therefore it must be true that

$$F_{(n)} > F_{(n-1)} > \dots > F_{(j+1)} > F_{(j)} \geq 1 - \lambda \quad (\text{A.1})$$

for if it were not true, then index $(j - 1)$ would have satisfied Ψ_{\max} and would have been an element of ζ_{\max} , which contradicts the initial supposition that it is not an element ζ_{\max} . Then by equation 7 in Horrace (2005), $\Pr\{CS\} = \Pr\{[n] \in \zeta_{\max} \geq 1 - \lambda\}$. Furthermore, ζ_{\max} has minimum cardinality, since replacing any index $(s) \in \zeta_{\max}$ with any index $(t) \notin \zeta_{\max}$ will clearly violate (A.1), while augmenting ζ_{\max} with any $(t) \notin \zeta_{\max}$ will satisfy (10), but the cardinality of ζ_{\max} will have been increased by one. Similar arguments can be made to prove the results for ζ_{\min} . \square

REFERENCES

- Aigner DJ, Lovell AK, Schmidt P. 1977. Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics* **6**: 21–37.
- Atkinson SE, Dorfman JH. 2005. Multiple comparisons with the best: Bayesian precision measures of efficiency rankings. *Journal of Productivity Analysis* **23**: 359–382.
- Battese GE, Coelli TJ. 1988. Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics* **38**: 387–399.
- Battese GE, Coelli TJ. 1992. Frontier production functions, technical efficiency, and panel data. *Journal of Productivity Analysis* **3**: 153–169.
- Battese GE, Coelli TJ. 1993. A stochastic frontier production incorporating a model for technical inefficiency effects. *Working Papers in Econometrics and Applied Statistics* No. 69, Department of Econometrics, University of New England, Armidale, Australia.
- Battese GE, Coelli TJ. 1995. A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics* **20**: 325–332.
- Battese GE, Coelli TJ, Colby TC. 1989. Estimation of frontier production functions and the efficiencies of Indian farms using panel data from ICRISAT's village level studies. *Journal of Quantitative Economics* **5**: 327–348.
- El-Gamal M, Grether D. 1995. Are people Bayesian? Uncovering behavior strategies. *Journal of American Statistical Association* **90**: 1137–1145.
- El-Gamal M, Grether D. 1999. Changing decision rules: uncovering behavioral strategies using estimation/classification (EC). In *Preferences, Beliefs, and Attributes in Decision Making*, Machina M, and Munier B (eds). Kluwer: New York; 239–251.
- El-Gamal M, Inanoglu H. 2005. Inefficiency and heterogeneity in Turkish banking: 1990–2000. *Journal of Applied Econometrics* **20**: 641–664.
- Felthoven RG. 2002. Effects of the American Fisheries Act. *Marine Resource Economics* **17**: 181–206.
- Fernandez C, Koop G, Steel MFJ. 2002. Multiple output production with undesirable outputs: an application to nitrogen surplus in agriculture. *Journal of the American Statistical Association* **97**: 432–442.
- Gupta SS. 1965. On some multiple decision (selection and ranking) rules. *Technometrics* **7**: 225–245.
- Guyader O, Daures F, Fifas S. 2004. A bioeconomic analysis of the impact of decommissioning programs: application to a limited-entry French scallop fishery. *Marine Resource Economics* **19**: 225–242.
- Horrace WC. 2005. On ranking and selection from independent truncated normal distributions. *Journal of Econometrics* **126**: 335–354.
- Horrace WC, Schmidt P. 2000. Multiple comparisons with the best, with economic applications. *Journal of Applied Econometrics* **15**: 1–26.
- Horrace WC, Marchand JT, Smeeding TM. 2007. Ranking inequality: applications of multivariate subset selection. *Journal of Economic Inequality* **5**: (forthcoming).
- Kim Y, Schmidt P. 2000. A review and empirical comparison of Bayesian and classical approaches to inference on efficiency levels in stochastic frontier models with panel data. *Journal of Productivity Analysis* **14**: 91–118.

- Kirkley J, Squires D, Strand EI. 1998. Characterizing managerial skill and technical efficiency in a fishery. *Journal of Productivity Analysis* **9**: 145–160.
- Kirkley J, Morrison Paul CJ, Squires D. 2002. Capacity and capacity utilization in common-pool resource industries. *Environmental and Resource Economics* **22**: 71–97.
- Kirkley J, Morrison Paul CJ, Squires D. 2004. Deterministic and stochastic capacity estimation for fishery capacity reduction. *Marine Resource Economics* **19**: 271–294.
- Koop G, Osiewalski J, Steel MFJ. 1997. Bayesian efficiency analysis through individual effects: hospital cost frontiers. *Journal of Econometrics* **76**: 77–105.
- Meeusen W, van den Broeck J. 1977. Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review* **18**: 435–444.
- Pascoe S, Coglan L. 2002. The contribution of unmeasurable inputs to fisheries production: an analysis of technical efficiency of fishing vessels in the English Channel. *American Journal of Agricultural Economics* **84**: 585–597.
- Schnier KE, Anderson CM, Horrace WC. 2006. Estimating heterogeneous production in fisheries. Working paper, Department of Environmental and Natural Resource Economics, University of Rhode Island.
- Sharma KR, Leung PS. 1998. Technical efficiency of Hawaii's longline fishery. *Marine Resource Economics* **13**: 259–274.
- Tsionas EG. 2002. Stochastic frontier models with random coefficients. *Journal of Applied Econometrics* **17**: 127–147.