

**MARGINAL COMPARISONS WITH THE BEST  
AND THE EFFICIENCY MEASUREMENT PROBLEM**

**Yangseon Kim  
East-West Center**

**Peter Schmidt  
Michigan State University**

**December, 1998**

**This version: December, 2006**

The second author acknowledges support under a Marie Curie Transfer of Knowledge Fellowship of the European Community under contract MTKD-CT-014288.

## 1. INTRODUCTION

Suppose that we have data on each of a set of  $N$  populations, indexed by a parameter  $\theta_i$ ,  $i = 1, 2, \dots, N$ . The parameterization is such that a larger value of  $\theta_i$  is "better" than a smaller value.

Suppose that we order the  $\theta_i$  as follows:

$$(1) \quad \theta_{(1)} \leq \theta_{(2)} \leq \dots \leq \theta_{(N)},$$

so that population (N) is best. The identity of the best population is *not* assumed to be known, which is the challenging aspect of the problem. Let  $\theta \equiv (\theta_1, \theta_2, \dots, \theta_N)'$  be the (unordered) vector of parameters. We presume that we have data on each of the populations and correspondingly there is an estimate  $\hat{\theta}$  of  $\theta$ . Based on this estimate we wish to say which populations might be best, and to construct confidence intervals for the differences  $\theta_{(N)} - \theta_i$ , which measure the amount by which a given population differs from the best. Thus we are interested in the problem of *comparisons with the best population*.

One case where comparisons with the best arise naturally is the measurement of productive efficiency. Following Schmidt and Sickles (1984) and much subsequent literature, we consider a fixed effect panel data production function of the form:

$$(2) \quad y_{it} = \theta_i + x_{it}'\beta + v_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

where  $i$  indexes firms and  $t$  indexes time. The production frontier is defined by the best firm, which is the one with the largest value of  $\theta_i$ , and one measure of the technical inefficiency of a given firm is the difference between the best firm's intercept and the given firm's intercept. We will consider this case in more detail in Section 4 below. However, there are many other settings in which comparison with the best may be appropriate. One example is a drug trial, in which  $\theta_i$  is the mean survival time (or some similar parameter where a larger value is better) given treatment with drug  $i$ . Suppose there is an existing standard drug plus a number of new possible drugs. We might be interested in comparing any one of the new drugs to the existing standard drug. However, we might also be interested in comparing new drugs with the best drug in the set. For example, a comparison with the best might be appropriate if one of the new drugs is very expensive and therefore is justified only if it is clearly best. Hsu (1996) gives a number of other

examples, such as the effectiveness of insect traps and the comparison of SAT scores across academic units.

One solution to the comparison with the best problem is given by the technique of *multiple comparisons with the best*, or MCB. MCB constructs a set  $S$  of possibly best populations, and a set of intervals  $(L_i, U_i)$ , such that:

$$(3) \quad P[ (N) \in S \text{ and } L_i \leq \theta_{(N)} - \theta_i \leq U_i \text{ for all } i ] \geq 1-\alpha ,$$

where  $1-\alpha$  is a chosen confidence level (e.g. 0.95). Thus with a given confidence level we have a set of populations that includes the best, and joint confidence intervals for all differences from the best. MCB was developed by Hsu (1981, 1984) and Edwards and Hsu (1983). A general exposition can be found in Hochberg and Tamhane (1987), Hsu (1996) and Horrace and Schmidt (2000).

An alternative to the multiple confidence intervals in (3) is a marginal (i.e., univariate) confidence interval for  $\theta_{(N)} - \theta_i$ , for a single given value of  $i$ . In the efficiency measurement example, this would amount to a confidence interval for the technical inefficiency of a given firm, which is a natural and useful object of interest. Perhaps surprisingly, the construction of a marginal confidence interval for  $\theta_{(N)} - \theta_i$  is a previously unsolved problem. In this paper we show how to construct these marginal confidence intervals. More precisely, for a given value of  $i$ , we provide a set  $S$  and an interval  $(L_i^m, U_i^m)$  such that  $P[ (N) \in S \text{ and } L_i^m \leq \theta_{(N)} - \theta_i \leq U_i^m ] \geq 1-\alpha$ . The point is that marginal confidence intervals are natural to consider, and also that we would expect marginal confidence intervals to be narrower than joint ones.

Horrace and Schmidt (1996, 2000) have applied MCB to the efficiency measurement problem. In this paper we provide similar applications of marginal comparisons with the best.

## 2. MARGINAL COMPARISONS UNDER STANDARD ASSUMPTIONS

Throughout the paper we will maintain the following two assumptions. First, we have an estimate  $\hat{\theta}$  distributed as  $N(\theta, \sigma^2 C)$  with the  $N \times N$  matrix  $C$  known. Second, either  $\sigma^2$  is known, or

we have an estimate  $\hat{\sigma}^2$ , independent of  $\hat{\theta}$ , such that  $\hat{\sigma}^2/\sigma^2$  is distributed as  $\chi_v^2/v$ . In any applications we envision, there will be enough degrees of freedom ( $v$  will be large enough) that we can effectively take  $\sigma^2$  as known.

The above assumptions put us squarely in the finite sample setting, since we require that  $\hat{\theta}$  be unbiased and normally distributed. One could alternatively pursue an analysis in terms of asymptotics (as  $T$  increased with  $N$  fixed) in which  $\hat{\theta}$  was consistent and asymptotically normal, with known asymptotic variance matrix. However, the MCB literature typically does not pursue the details of this, and neither will this paper.

*Standard* MCB proceeds under the further assumption, which we will maintain in this section, that  $C = kI_N$  with the scalar  $k$  known. This assumption is usually motivated by discussion of the "balanced one way model" (e.g., Hsu (1996), p. 43) in which we have independent observations  $y_{it}$  ( $i = 1, \dots, N, t = 1, \dots, T$ ) distributed as  $N(\theta_i, \sigma^2)$ . In this case  $\hat{\theta}_i = \bar{y}_i$ , and so we have  $C = I_N$  and  $k = 1/T$ . This is also the case in the panel data regression model with fixed individual effects, if we treat the slope coefficients as known, as will be discussed below.

We define the following notation, which is more or less standard in the MCB literature.  $E(1/2)$  is the  $(N-1) \times (N-1)$  correlation matrix with all correlations equal to  $1/2$  (i.e., diagonal elements equal one, off-diagonal elements equal  $1/2$ ). Let  $z$  be a multivariate random variable distributed as student-t with dimension  $N-1$ , degrees of freedom  $v$ , and correlation matrix  $E(1/2)$ . Define  $d^*(\alpha)$  as the  $\alpha$ -level critical value of  $\max_{i=1, \dots, N-1} |z_i|$ ; i.e.,  $P[\max_i |z_i| \leq d^*(\alpha)] = 1-\alpha$ . Tabulations of  $d^*(\alpha)$  can be found in Hsu (1996) or Horrace (1998). Define  $h(\alpha) = d^*(\alpha)(2k\hat{\sigma}^2)^{1/2}$ , and define the set  $S(\alpha) = \{i \mid \hat{\theta}_i \geq \max_{j=1, \dots, N} \hat{\theta}_j - h(\alpha)\}$ . Define  $L_i$  and  $U_i$  as follows:

$$(4) \quad L_i = \max[0, \min_{j \in S(\alpha)} \hat{\theta}_j - \hat{\theta}_i - h(\alpha)], \quad U_i = \max[0, \max_{j \neq i} \hat{\theta}_j - \hat{\theta}_i + h(\alpha)]$$

Then MCB provides the statement (3) above, with  $S = S(\alpha)$ . See, e.g., Horrace and Schmidt (2000, p. 6).

Our marginal comparison with the best is given by the following theorem.

**THEOREM 1:** Let  $t^*(\alpha)$  be the two-sided  $\alpha$ -level critical value of the (univariate) student-t distribution with  $\nu$  degrees of freedom; i.e., if  $z$  is distributed as student-t with  $\nu$  degrees of freedom, then  $P[|z| \leq t^*(\alpha)] = 1-\alpha$ . Define  $g(\alpha) = t^*(\alpha)(2k\hat{\sigma}^2)^{1/2}$ . Define the set  $S(\alpha)$  as above. Define  $L_i^m$  and  $U_i^m$  as follows:

$$(5) \quad L_i^m = \max[0, \min_{j \in S(\alpha)} \hat{\theta}_j - \hat{\theta}_i - g(\alpha/2)], \quad U_i^m = \max[0, \max_{j \neq i} \hat{\theta}_j - \hat{\theta}_i + g(\alpha/2)].$$

Then

$$(6) \quad P[(N) \in S(\alpha) \text{ and } L_i^m \leq \theta_{(N)} - \theta_i \leq U_i^m] \geq 1-\alpha.$$

The proof is given in the Appendix. It is a relatively straightforward application of the Bonferroni inequality and the method of proof used to establish the MCB result..

We can note that the marginal comparison (6) uses the  $\alpha/2$ -level critical value of a univariate student-t while the multiple comparison (3) uses the  $\alpha$ -level critical value of an  $N-1$  dimensional student-t. There is no general inequality between these, but for commonly chosen values of  $\alpha$  (e.g. 0.05 or 0.10) the marginal intervals are narrower than the multiple intervals except when  $N$  is quite small. For example, for  $\alpha = 0.05$  and  $\nu = \infty$ , the univariate  $\alpha/2$ -level critical value is 2.24, while the  $N-1$  variate  $\alpha$ -level critical value is 2.21 for  $N=3$ , 2.35 for  $N=4$ , 2.44 for  $N=5$ , 2.69 for  $N=10$ , 3.01 for  $N=30$ , 3.30 for  $N=100$ , and so forth.

We may also wish to consider one-sided confidence intervals. One of the possible motivations for doing so is the following. In many applications, the lower bound for  $\theta_{(N)} - \theta_i$  turns out to be zero for many observations, because the set  $S$  of possibly best populations is large. We might choose to forgo the calculation of a lower bound, in which case a tighter upper bound is possible. This result is given in the following theorem.

**THEOREM 2:** Let  $g(\alpha)$  be defined as in the statement of Theorem 1. Then the following are true:

$$(7) \quad P[(N) \in S(\alpha) \text{ and } \theta_{(N)} - \theta_i \leq \max[0, \max_{j \neq i} \hat{\theta}_j - \hat{\theta}_i + g(\alpha)]] \geq 1-\alpha$$

$$(8) \quad P[\theta_{(N)} - \theta_i \leq \max[0, \max_{j \neq i} \hat{\theta}_j - \hat{\theta}_i + g(2\alpha)]] \geq 1-\alpha.$$

The proof is given in the Appendix, but we can note the following. Comparing (7) to (6), the fact that we make only one statement instead of two allows us to use the  $\alpha/2$ -level *one-sided* univariate student-t critical value, which is the same as the  $\alpha$ -level two-sided critical value, instead of the  $\alpha/2$ -level two-sided critical value. (For example, for  $\alpha = 0.05$  and  $v = \infty$ , we use 1.96 instead of 2.24.) Considering (8), we note that the upper bound does not require the definition of the possibly best set  $S(\alpha)$ . If we do not wish to consider  $S(\alpha)$ , we can devote the full confidence level  $1-\alpha$  to the upper bound, and we can use the  $\alpha$ -level one-sided critical value of student-t, which is the same as the  $2\alpha$ -level two-sided critical value. (Thus, for example, with  $\alpha = 0.05$  and  $v = \infty$ , we can now use the critical value 1.64 instead of 1.96 or 2.24.) As a result we get more a precise upper bound.

### 3. MARGINAL COMPARISONS WITH GENERAL COVARIANCE STRUCTURE

The previous section considered the commonly-assumed special case that the covariance matrix of  $\hat{\theta}$  is proportional to an identity matrix. In this section we consider the general case that  $\hat{\theta}$  is distributed as  $N(\theta, \sigma^2 C)$  with  $C$  known but unrestricted. This arises in, among other cases, the panel data regression model with nontrivial regressors.

We first need to define a little notation. For a given value of  $i$ , define  $\delta_i$  as the  $(N-1) \times 1$  vector whose typical element is of the form  $\theta_j - \theta_i$ , for  $j = 1, \dots, N$ ,  $i \neq j$ . Formally  $\delta_i = D_i \theta$  where  $D_i$  is an  $(N-1) \times N$  differencing matrix. The covariance matrix of  $\hat{\delta}_i$  is  $\sigma^2 B_i$ , where  $B_i = D_i C D_i'$ . Let  $R_i$  be the corresponding correlation matrix. In the special case that  $C$  is proportional to identity,  $R_i = E(1/2)$ , as discussed in the previous section. In the general case,  $R_i$  will depend on  $i$  and has no special structure, but it is easily calculated. Define  $d_i^*(\alpha)$  as the two-sided  $\alpha$ -level critical value of the multivariate student-t distribution with dimension  $N-1$ , degrees of freedom  $v$ , and correlation matrix  $R_i$ . This critical value will typically depend on  $i$  and will generally need to be calculated numerically (e.g. by a simulation), since tabulation is impossible except in special cases.

Now define  $\hat{\sigma}_{ij}^2 = \hat{\sigma}_{ji}^2 = [\hat{\sigma}^2 (C_{ii} + C_{jj} - 2C_{ij})]$ ,  $h_{ji}(\alpha) = d_j^*(\alpha) \hat{\sigma}_{ij}$ ,  $L_i^j = \hat{\theta}_j - \hat{\theta}_i - h_{ji}(\alpha)$ ,  $U_i^j = \hat{\theta}_j - \hat{\theta}_i + h_{ji}(\alpha)$ , the possibly best set  $S(\alpha) = \{i \mid U_j^i \geq 0 \forall j \neq i\} = \{i \mid \hat{\theta}_i \geq \hat{\theta}_j - h_{ij}(\alpha) \forall j \neq i\}$ , and the lower and upper bounds  $L_i = \max[0, \min_{j \in S(\alpha)} L_i^j]$  and  $U_i = \max[0, \max_{j \neq i} U_i^j]$ . Then MCB provides the statement (3) above, with  $S = S(\alpha)$ . See, e.g., Horrace and Schmidt (2000, p. 10).

We now provide the corresponding marginal comparison result.

**THEOREM 3:** Define the set  $S(\alpha)$  as above, and let  $t^*(\alpha)$  be the two-sided  $\alpha$ -level critical value of the univariate student-t distribution. Define  $g_{ij}(\alpha) = t^*(\alpha) \hat{\sigma}_{ij}$ . Define  $L_i^m$  and  $U_i^m$  as follows:

$$(9) \quad L_i^m = \max[0, \min_{j \in S(\alpha)} (\hat{\theta}_j - \hat{\theta}_i - g_{ij}(\alpha/2))], \quad U_i^m = \max[0, \max_{j \neq i} (\hat{\theta}_j - \hat{\theta}_i + g_{ij}(\alpha/2))].$$

Then

$$(10) \quad P[(N) \in S(\alpha) \text{ and } L_i^m \leq \theta_{(N)} - \theta_i \leq U_i^m] \geq 1 - \alpha.$$

The proof is similar to the proof of Theorem 1 and is therefore omitted.

As in the standard case, we may also consider one-sided confidence intervals. The following Theorem (also presented without proof) is the result corresponding to Theorem 2.

**THEOREM 4:** The following are true:

$$(11) \quad P[(N) \in S(\alpha) \text{ and } \theta_{(N)} - \theta_i \leq \max[0, \max_{j \neq i} (\hat{\theta}_j - \hat{\theta}_i + g_{ij}(\alpha))] \geq 1 - \alpha$$

$$(12) \quad P[\theta_{(N)} - \theta_i \leq \max[0, \max_{j \neq i} (\hat{\theta}_j - \hat{\theta}_i + g_{ij}(2\alpha))] \geq 1 - \alpha .$$

As in the previous section, one possible motivation for one-sided confidence intervals is that they yield more precise upper bounds. However, in the case of general covariance structure the one-sided intervals given in equation (12) also offer considerable computational advantages, because they do not require the calculation of the possibly best set  $S(\alpha)$ . The calculation of  $S(\alpha)$  requires the  $N$  critical values  $d_i^*(\alpha)$ ,  $i = 1, \dots, N$ , each of which is from an  $N-1$  dimensional student-t distribution and is generally calculable only numerically (via simulation). Especially when  $N$  is large, this is a very complicated and time-consuming set of calculations.

#### 4. THE EFFICIENCY MEASUREMENT PROBLEM

In this section we discuss the application of our marginal comparisons with the best to the efficiency measurement problem. Horrace and Schmidt (1996, 2000) applied MCB to this problem, and we wish to compare the marginal and multiple confidence intervals. Here we will give a very brief discussion of the problem. More detail can be found in Horrace and Schmidt (2000).

We begin with the fixed-effect panel data regression model:

$$(13) \quad y_{it} = \theta_i + x_{it}'\beta + v_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T.$$

This is the same as equation (2) above. For purposes of exposition we assume a "balanced" panel (T is the same for all i). We assume that the  $v_{it}$  are iid  $N(0, \sigma_v^2)$  and we treat the explanatory variables  $x_{it}$  as fixed. The parameters of interest are the intercepts  $\theta_i$ . In our applications the estimator of  $\beta$ , say  $\hat{\beta}$ , is the fixed-effect or "within" estimator obtained by least squares of  $\check{y}_{it}$  on  $\check{x}_{it}$ , where  $\check{y}_{it} = y_{it} - \bar{y}_i$  and where  $\bar{y}_i$  is the mean of the T observations on  $y_{it}$  for firm i; and similarly for  $\check{x}_{it}$  and  $\bar{x}_i$ . Then we obtain an estimate of  $\theta_i$  as  $\hat{\theta}_i = \bar{y}_i - \bar{x}_i' \hat{\beta}$ . These are standard results from the panel data literature. More detail in the present setting can be found in Schmidt and Sickles (1984).

In the context of the efficiency measurement problem,  $y$  is the logarithm of output and  $x$  is a vector of functions of inputs into the productive process. Larger  $\theta_i$  is better because it corresponds to more output for the same inputs. We define  $u_i = \theta_{(N)} - \theta_i$ , where  $\theta_{(N)}$  is the largest of the N  $\theta_i$ 's, and the *technical efficiency* of firm i is typically defined as  $TE_i = \exp(-u_i)$ . Since  $u_i \geq 0$ ,  $0 \leq TE_i \leq 1$ . MCB and our marginal comparisons with the best procedures will provide confidence intervals for  $u_i$ , and these are easily converted into confidence intervals for  $TE_i$ . In particular, if (with a given probability)  $L \leq u_i \leq U$ , then  $\exp(-U) \leq TE_i \leq \exp(-L)$ , so that lower bounds for  $u_i$  convert to upper bounds for  $TE_i$  and conversely.

Let  $\theta = (\theta_1, \theta_2, \dots, \theta_N)'$ . Under our assumptions the covariance matrix of  $\hat{\theta}$  is not proportional to an identity matrix and so we should allow for a general covariance structure.

Specifically,  $V(\hat{\theta}) = (\sigma_v^2/T)I_N + \bar{X} V(\hat{\beta}) \bar{X}'$ , where  $V(\hat{\beta})$  is the covariance matrix of  $\hat{\beta}$  and  $\bar{X}$  is the matrix whose  $i^{\text{th}}$  row is  $\bar{x}_i'$ . Furthermore  $V(\hat{\beta}) = \sigma_v^2 (\ddot{X}\ddot{X})^{-1}$  where  $\ddot{X}$  is the matrix whose typical row is  $\ddot{x}_i'$ . Therefore

$$(14) \quad V(\hat{\theta}) = \sigma_v^2 C, \text{ where } C = (1/T)I_N + \bar{X}(\ddot{X}\ddot{X})^{-1} \bar{X}' .$$

Finally, evaluating this requires an estimate of  $\sigma_v^2$ , for which the standard choice is  $\hat{\sigma}_v^2 = \text{SSE}/[N(T-1)]$ , where SSE is the unexplained sum of squares in the “within” regression (regression of  $\ddot{y}$  on  $\ddot{X}$ ).

The methods of section 2 ("standard" MCB or marginal comparisons) would apply if  $\beta$  were known, since then the term arising from  $V(\hat{\beta})$  disappears and  $C = (1/T)I_N$ . Therefore standard MCB can be viewed as applying approximately if the proportion of  $V(\hat{\beta})$  due to the variance of  $\hat{\beta}$  is small. This may generally be so when  $N$  is large relative to  $T$ , as discussed in Horrace and Schmidt (2000). However, in this paper we present only the results that allow for the general covariance structure.

## 5. EMPIRICAL RESULTS

### 5.1 Indonesian Rice Farms

We first analyze the data of Erwidodo (1990), which contain information on  $N = 171$  rice farms for  $T = 6$  growing seasons. Output is the physical quantity of rice grown, and inputs included are seed, fertilizer, labor, land, and some dummy variables. A Cobb-Douglas (linear in logarithms) functional form is assumed. More detail can be found in Erwidodo (1990) or Horrace and Schmidt (1996).

Our results are given in Table 1. We choose  $\alpha = 0.10$  (hence 90% confidence intervals). We give results for the three most efficient (best) firms, the 75<sup>th</sup> percentile, 50<sup>th</sup> percentile (median) and 25<sup>th</sup> percentile firms, and the two least efficient (worst) firms. For each firm, we present the value of  $\hat{\theta}$ ; the estimate of technical efficiency; the confidence intervals corresponding to the marginal and multiple comparisons with the best; and the corresponding

one-sided marginal and one-sided multiple comparisons with the best. The one-sided marginal comparisons correspond to equation (12) of section 3.

In these data we do not estimate the intercepts  $\theta_i$  very precisely. This occurs because we have only six observations per firm and because the value of  $\sigma_v^2$  turns out to be large relative to the variation in the  $\theta_i$ . (Our estimate of  $\sigma_v^2$  is 0.108 and the sample variance of the  $\hat{\theta}_i$  equals 0.031.)

Correspondingly, our confidence intervals are rather wide. In fact, they are wide enough to suggest that the efficiency measurement exercise has more or less failed to distinguish efficient and inefficient firms. The possibly best set S contains 92 of the 171 firms, and the other 79 firms are sufficiently close to being in the possibly best set that the upper bounds for  $TE_i$  equal one for all firms, even for the two-sided intervals. (For the one-sided intervals, the upper bound is automatically one, but this is an identity, not a data-determined outcome.)

The marginal comparison intervals are considerably shorter than the MCB intervals, as they should be. They use the 5% critical value of the univariate student-t distribution, 1.96, while the multiple intervals use 10% critical values of the 170-dimensional student-t distribution, which vary a little over comparison populations (i.e.  $d_i^*$  above depends on  $i$ ) but equal 3.18 on average.

The greater precision of the marginal as opposed to multiple confidence intervals is most noticeable for the more efficient firms. For example, for the most efficient firm compare the marginal interval of [0.74,1] to the multiple interval of [0.58,1]. The extra width of the MCB intervals is the price one has to pay for making a multiple statement, of course (the so-called "multiplicity effect").

## **5.2 Texas Utilities**

We next analyze the data of Kumbhakar (1994). We observe the output of electrical power and the use of labor, capital and fuel as inputs by  $N = 10$  major privately-owned electric utilities in Texas. There are  $T = 18$  yearly observations. We estimate the production function (Kumbhakar estimated the cost function) in Cobb-Douglas form.

The results are given in Table 2. The format of this Table is the same as that of Table 1 except that now we are able to show results for all of the firms. The confidence intervals (both marginal and multiple) are much narrower for this data set than for the previous one. We are now able to make statements about efficiencies that are precise enough to be meaningful. For example, only two observations are in the possibly best set, and the confidence interval for the efficiency of the (apparently) most efficient firm is [0.95,1]. This occurs primarily because  $\sigma_v^2$  is smaller and T is larger here than in the previous data set. (Our estimate of  $\sigma_v^2$  is 0.003 and the sample variance of the  $\hat{\theta}_i$  equals 0.007. Also now we have T=18 rather than T=6. As a result the  $\theta_i$  are estimated more precisely.) The marginal intervals are narrower than the multiple intervals but the difference in width is not as large as it was in the previous data set, because the multiplicity effect is weaker with N = 10 than with N = 171. Numerically, the marginal intervals use the univariate student-t 5% critical value of 1.96, while the multiple intervals use 10% nine-variate student-t critical values, which are on average equal to 2.38.

### 5.3 Some Specification Issues

The empirical analyses above, and many other empirical analyses in the efficiency measurement literature, rely on some strong assumptions, and in this section we consider how to test and (if necessary) relax some of these assumptions. Specifically, we will consider two issues. The first issue is the assumption of *strict exogeneity*, which is the assumption that  $v_{it}$  is uncorrelated with  $x_{is}$  for all t and s. This assumption is necessary for the within estimator to be consistent. It is implied by our assumption above that the  $x_{it}$  can be treated as fixed; that is, that they are independent of the errors. It is a strong assumption because it rules out the possibility that current  $v$  can affect future  $x$ , that is, the possibility that current noise could affect future input choices. The second issue is the assumption of *white noise* errors, that is, the assumption that  $\text{var}(v_{it})$  is constant and that  $\text{cov}(v_{it}, v_{is}) = 0$  for  $t \neq s$ . This assumption is necessary for the form of C given in equation (14) to be correct.

We first consider the strict exogeneity assumption. This can be tested in a large number of ways. The simplest is to add  $x_{i,t+1}$  to the regression (13), and then to do an F-test of the significance of these variables after fixed effects estimation. When we do this, we accept the null hypothesis of strict exogeneity for the Indonesian rice farm data ( $F_{8,667} = 1.65$  vs. the 95% critical value of 1.94), but we reject the null hypothesis for the Texas utility data ( $F_{3,164} = 3.13$  vs. the 95% critical value of 2.60). These tests are conducted under the assumption that the errors are white noise, but the use of a robust variance matrix for the estimated regression coefficients did not change the results of the tests

For the Texas utility data, we therefore proceed to estimation by instrumental variables (IV), whose consistency requires only weak exogeneity ( $v_{it}$  is uncorrelated with  $x_{is}$  for  $s \leq t$ ). Specifically, we take first differences of the regression (equation (13)), and then estimate by IV, where the instrumental variables are the lagged x terms:  $x_{i,t-1}, x_{i,t-2}, \dots, x_{i,t-p}$ . We did this for  $p = 1, 2$  and  $3$  and the results were not very different. Table 3 gives the confidence intervals for the case of  $p = 3$ . The results are relatively similar to those in Table 2. (For example, for firm number 8, which is one of the median efficiency firms, we now have a MargCB interval of  $[0.756, 0.972]$  instead of  $[0.764, 0.964]$ .) We conclude that, even though we have rejected the strict exogeneity assumption, relaxing it does not affect the results very much.

Next we test the white noise assumption. We use the minimum MCS test of Arellano (1990, p. 130), for his case (ii), white noise errors. For the Indonesian rice farm data, we reject the white noise hypothesis ( $\chi_{14}^2 = 166.50$  vs. the 95% critical value of 23.7). For the Texas utility data, the test statistic is not defined because  $N$  is too small relative to  $T$ . (The estimated variance matrix of the first differenced errors is singular.) So we based the test just on the leading 3 by 3 submatrix of the variance matrix of the first differenced errors. This test accepted the white noise hypothesis ( $\chi_5^2 = 6.21$  vs. the 95% critical value of 11.1).

For the Indonesian rice farm data, we therefore would like to construct confidence intervals that are robust to the failure of the white noise assumption. This is a non-trivial task. Although it is easy to construct a robust variance matrix estimate for the fixed-effects estimates of the regression coefficients, there is no robust variance matrix for the vector  $\hat{\theta}$  of estimated individual effects. However, we can estimate an *upper bound* for  $V(\hat{\theta})$ . Technical details can be found in the Appendix. Using this upper bound, we get conservative but valid confidence intervals. These are given in Table 4. They are modestly wider than the corresponding intervals in Table 1. (For example, for firm 15, the median efficiency firm, we now have a MargCB interval of [0.339, 1] instead of [0.379, 1].) So using a conservative robust variance matrix for  $\hat{\theta}$  makes a little difference but not much.

## 6. CONCLUDING REMARKS

In this paper we have considered the general problem of creating confidence intervals for measures of the difference between a given population and the best population. More precisely, population  $i$  is characterized by a parameter  $\theta_i$ , and we wish to construct a confidence interval for the difference  $\theta_{(N)} - \theta_i$ , where  $\theta_{(N)} = \max_{j=1, \dots, N} \theta_j$ . This is a challenging problem because we do not know which population is best. One solution is given by MCB, which provides the complete set of  $N$  such confidence intervals, all of which hold simultaneously with at least a specified confidence level. Perhaps surprisingly, the seemingly simpler problem of providing a confidence interval for a single difference  $\theta_{(N)} - \theta_i$  had not previously been solved. In this paper we provide these confidence intervals, and refer to them as *marginal* comparisons with the best.

Whether one prefers multiple or marginal comparisons will no doubt depend on the context. For an example of the arguments in favor of multiple comparisons, see Hsu (1996, p. 7). However, in some cases a marginal comparison may be natural. It seems reasonable to be able to perform either type of inference, just as one may wish to be able to test a set of hypotheses either individually or jointly.

In the context of the efficiency measurement problem, marginal comparisons correspond to the construction of the confidence interval for a given firm's technical efficiency level, and this is indeed a natural thing to consider. For example, models that assume a distribution for  $u_i$  yield marginal confidence intervals, constructed in somewhat more straightforward ways than here. See Horrace and Schmidt (1996) or Koop *et al.* (1997) for some examples. A marginal comparison with the best is directly comparable, and provides evidence on the gain in precision from assuming a distribution for  $u_i$ . These comparisons are harder when MCB is used because the multiplicity effect and the effect of assuming or not assuming a distribution become confounded.

There is a literature on Bayesian methods for the efficiency measurement exercise, to which a good introduction is Koop *et al.* (1997). Bayesian methods also allow comparisons with the best, and do so in a methodologically straightforward and consistent way, though at the cost of considerable computational complexity. We are not aware of Bayesian comparisons with the best outside of the efficiency measurement literature, but there is no reason why the methodology could not be applied more widely.

## APPENDIX A

### Proof of Theorem 1

As in the text, suppose that  $z$  is a multivariate random variable distributed as student-t with dimension  $N-1$ , degrees of freedom  $\nu$ , and correlation matrix  $E^{(1/2)}$ . Define  $d_1^*(\alpha)$  as the  $\alpha$ -level critical value of  $\max_{i=1, \dots, N-1} z_i$ ; i.e.,  $P[\max_i z_i \leq d_1^*(\alpha)] = 1-\alpha$ . Note that  $d_1^*(\alpha)$  is the one-sided critical value corresponding to the two-sided critical value  $d^*(\alpha)$  used in MCB, and that  $d_1^*(\alpha/2) = d^*(\alpha)$ . Similarly define  $h_1(\alpha) = d_1^*(\alpha)(2k\hat{\sigma}^2)^{1/2}$  and note that  $h_1(\alpha/2) = h(\alpha)$ , with  $h(\alpha) = d^*(\alpha)(2k\hat{\sigma}^2)^{1/2}$ , as used in MCB.

Consider the event  $E_1(\alpha) = \{\theta_{(N)} - \theta_j \leq \hat{\theta}_{(N)} - \hat{\theta}_j + h_1(\alpha) \forall j \neq (N)\}$ . This is the one-sided multiple comparisons with a control (MCC) event, with  $(N)$  as control, and is constructed so that  $P[E_1(\alpha)] = 1-\alpha$ . See Dunnett (1955, 1964) or the discussion in Hsu (1996, chapter 3). The event  $E_1(\alpha)$  implies the event  $\{(N) \in S_1(\alpha)\}$ , where  $S_1(\alpha)$  is the set of indices  $S_1(\alpha) = \{i \mid \hat{\theta}_i \geq \max_{j=1, \dots, N} \hat{\theta}_j - h_1(\alpha)\}$ . Note that  $S_1(\alpha/2) = S(\alpha) \subseteq S(\alpha/2)$ . Therefore

$$(A1) \quad P[(N) \in S_1(\alpha/2)] = P[(N) \in S(\alpha)] \geq 1-\alpha/2,$$

a standard result of the "ranking and selection" literature; e.g., see Gupta (1965).

Now pick a value of  $i$  ( $= 1, \dots, N$ ), and consider the event  $A_i(\alpha) = \{\hat{\theta}_{(N)} - \hat{\theta}_i - g(\alpha) \leq \theta_{(N)} - \theta_i \leq \hat{\theta}_{(N)} - \hat{\theta}_i + g(\alpha)\}$ , where  $g(\alpha)$  was defined in the statement of the Theorem. Note that  $g(\alpha)$  was constructed so that  $P[A_i(\alpha)] = 1-\alpha$ . By the Bonferroni inequality, it follows from (A1) and  $P[A_i(\alpha/2)] = 1-\alpha/2$  that

$$(A2) \quad P[(N) \in S(\alpha) \text{ and } A_i(\alpha/2)] \geq 1-\alpha.$$

This inequality is not immediately useful because it is not in terms of observable quantities, since  $(N)$  is unknown. So, we need to show that the event  $\{(N) \in S(\alpha) \text{ and } A_i(\alpha/2)\}$  implies the marginal comparison event given in (6) of the main text. Consider first the lower bound. The event whose probability is given in (A2) implies that

$$(A3) \quad \min_{j \in S(\alpha)} \hat{\theta}_j - \hat{\theta}_i - g(\alpha/2) \leq \hat{\theta}_{(N)} - \hat{\theta}_i - g(\alpha/2) \leq \theta_{(N)} - \theta_i.$$

Also  $0 \leq \theta_{(N)} - \theta_i$ . Thus the event whose probability is given in (A2) implies the lower bound

$$(A4) \quad \max[0, \min_{j \in S(\alpha)} \hat{\theta}_j - \hat{\theta}_i - g(\alpha/2)] \leq \theta_{(N)} - \theta_i.$$

The treatment of the upper bound is similar. If  $(N) = i$ , then  $\theta_{(N)} - \theta_i = 0$ . If  $(N) \neq i$ , the event in (A2) implies

$$(A5) \quad \theta_{(N)} - \theta_i \leq \max_{j \neq i} \hat{\theta}_j - \hat{\theta}_i + g(\alpha/2).$$

Therefore we have the upper bound

$$(A6) \quad \theta_{(N)} - \theta_i \leq \max[0, \max_{j \neq i} \hat{\theta}_j - \hat{\theta}_i + g(\alpha/2)].$$

Finally, since the event  $\{(N) \in S(\alpha)\}$  and the bounds (A4) and (A6) are implied by the event in (A2), they hold with at least the probability of that event; that is, with a probability no smaller than  $1-\alpha$ .

### Proof of Theorem 2

As in the proof of Theorem 6, we have  $P[(N) \in S(\alpha)] \geq 1-\alpha/2$ . Now we also have  $P[\theta_{(N)} - \theta_i \leq \hat{\theta}_{(N)} - \hat{\theta}_i + g(\alpha)] = 1-\alpha/2$ , since  $g(\alpha)$  is based on the  $\alpha$ -level two-sided student-t critical value, or equivalently the  $\alpha/2$ -level one-sided critical value. Thus the Bonferroni inequality implies that  $P[(N) \in S(\alpha) \text{ and } \theta_{(N)} - \theta_i \leq \hat{\theta}_{(N)} - \hat{\theta}_i + g(\alpha)] \geq 1-\alpha$ . Then the same logic as was used in the discussion leading up to (A6) yields the result in (7).

To establish equation (8), we note that the upper bound does not require the definition of the possibly best set  $S(\alpha)$ . We simply start with the statement:  $P[\theta_{(N)} - \theta_i \leq \hat{\theta}_{(N)} - \hat{\theta}_i + g(2\alpha)] = 1-\alpha$ , which follows from the fact that the two-sided  $2\alpha$ -level critical value in  $g(2\alpha)$  is the same as the  $\alpha$ -level one-sided critical value. Then we again apply the same logic as was used in the discussion leading up to (A6) to obtain (8).

### Robust $V(\hat{\theta})$

We still treat the exogenous variables as fixed, but now we allow the  $T$ -dimensional errors  $v_i$  to have an unrestricted variance matrix  $\Sigma$ . We avoid any assumptions about the form of  $\Sigma$ , other than that it is positive definite, and in order to do so we treat  $T$  as given (fixed). Thus any asymptotic arguments, such as consistency, need to be as  $N \rightarrow \infty$  with  $T$  fixed. However, the arguments below about what can and cannot be estimated do not really require asymptotics.

It is well known that under these assumptions we can construct a robust variance matrix for the fixed effects estimates of the regressions coefficients ( $\hat{\beta}$ ). It does not seem to be realized in the literature that this is not possible for the vector of estimated individual effects ( $\hat{\theta}$ ). To understand this point, we note that  $\hat{\theta} - \theta = \bar{v} - \bar{X}(\hat{\beta} - \beta)$ , and therefore

$$(A7) \quad V(\hat{\theta}) = A + B + C + C'$$

where

$$(A8a) \quad A = V(\bar{v}) = a \cdot I_N, \text{ where } a = e' \Sigma e / T^2 \text{ and where } e \text{ is a vector (of dimension } T) \text{ of ones;}$$

$$(A8b) \quad B = V(\bar{X}'(\hat{\beta} - \beta)) = \bar{X}' V(\hat{\beta}) \bar{X}, \text{ and } V(\hat{\beta}) = (\ddot{X} \ddot{X})^{-1} \sum_i \ddot{X}'_i \Sigma \ddot{X}_i (\ddot{X} \ddot{X})^{-1};$$

$$(A8c) \quad C = -E \bar{X}'(\hat{\beta} - \beta) \bar{v}' = -\bar{X}' (\ddot{X} \ddot{X})^{-1} X' [I_N \otimes Q \Sigma e] \text{ where } Q \text{ is the differencing matrix of dimension } T \times T.$$

In (A8b), we can write  $\ddot{X}'_i \Sigma \ddot{X}_i = X'_i Q \Sigma Q X_i$ . Thus, apart from observables, in order to evaluate A, B and C we need the three quantities  $e' \Sigma e$ ,  $Q \Sigma Q$  and  $Q \Sigma e$ . From Kiefer (1980), it is known that we can estimate  $Q \Sigma Q$  consistently; for example, a consistent estimate of  $Q \Sigma Q$  is  $\sum_i \hat{v}_i \hat{v}'_i / N$  where  $\hat{v}_i = \ddot{y}_i - \ddot{x}'_i \hat{\beta}$  are the within residuals. From Kiefer it is also known that we cannot estimate  $\Sigma$  consistently. Since knowledge of  $\Sigma$  is equivalent to knowledge of  $e' \Sigma e$ ,  $Q \Sigma Q$  and  $Q \Sigma e$ , it follows that it must be impossible to estimate  $e' \Sigma e$  and/or  $Q \Sigma e$ . Thus in equation (A7), we can evaluate the term B but one or more of the terms A and/or C cannot be evaluated.

It turns out that we can estimate  $Q \Sigma e$  consistently. Define  $\hat{v}_i = y_i - x'_i \hat{\beta}$  which are the residuals in the levels equation based on the within estimate  $\hat{\beta}$ . It is easy to see that  $\sum_i \hat{v}_i \hat{v}'_i / N$  is a consistent estimate of  $\Sigma + \sigma_\theta^2 e e'$ , where  $\sigma_\theta^2 = \lim_{N \rightarrow \infty} \sum_i (\theta_i - \bar{\theta})^2$  is interpreted as the variance of the  $\theta_i$ . It follows that  $Q(\sum_i \hat{v}_i \hat{v}'_i / N) e$  is a consistent estimate of  $Q(\Sigma + \sigma_\theta^2 e e') e = Q \Sigma e$ , where the last equality is due to the fact that  $Q e = 0$ . Thus term C in equation (A7) can also be evaluated.

This leaves the quantity  $a = e' \Sigma e / T^2$  as the thing that cannot be estimated consistently from the within regression, and the term A in equation (A7) as the term that cannot be evaluated.

Fundamentally this is because A is the variance matrix of  $\bar{v}$ , and  $\bar{v}$  is orthogonal to the within regression. However, we can find (estimate) an upper bound for the quantity. Specifically,  $e'(\sum_i \hat{v}_i \hat{v}_i' / N)e / T^2$  is a consistent estimate of  $e'(\Sigma + \sigma_\theta^2 ee')e / T^2 = e'\Sigma e / T^2 + \sigma_\theta^2 > e'\Sigma e / T^2$ .

Thus, using a consistent estimate of the terms B and C, and using a consistent estimate of an upper bound for term A, we can find a consistent estimate for an upper bound for  $V(\hat{\theta})$ . Using this upper bound leads to conservative but valid confidence intervals.

**Table 1**

90% Confidence Intervals for Technical Efficiency  
Indonesian Rice Farms

Firm No.	$\hat{\theta}_i$	TE <sub>i</sub>	MargCB		MCB		1-Sided MargCB		1-Sided MCB	
			LB	UB	LB	UB	LB	UB	LB	UB
164	5.556	1	0.737	1	0.583	1	0.840	1	0.615	1
118	5.486	0.932	0.642	1	0.508	1	0.730	1	0.534	1
163	5.484	0.930	0.643	1	0.509	1	0.729	1	0.535	1
:	:	:	:	:	:	:	:	:	:	:
31	5.072	0.616	0.421	1	0.328	1	0.478	1	0.346	1
:	:	:	:	:	:	:	:	:	:	:
15	4.966	0.554	0.379	1	0.300	1	0.432	1	0.315	1
:	:	:	:	:	:	:	:	:	:	:
16	4.859	0.498	0.340	1	0.266	1	0.387	1	0.280	1
:	:	:	:	:	:	:	:	:	:	:
117	4.586	0.379	0.259	0.974	0.203	1	0.295	1	0.214	1
45	4.550	0.365	0.250	0.940	0.197	1	0.285	1	0.208	1

**Table 2**

90% Confidence Intervals for Technical Efficiency  
Texas Utilities

Firm No.	$\hat{\theta}_i$	TE <sub>i</sub>	MargCB		MCB		1-Sided MargCB		1-Sided MCB	
			LB	UB	LB	UB	LB	UB	LB	UB
5	-4.995	1	0.960	1	0.945	1	1	1	0.967	1
3	-5.083	0.916	0.806	1	0.797	1	0.842	1	0.815	1
10	-5.145	0.861	0.773	0.979	0.765	0.984	0.802	1	0.780	1
1	-5.176	0.835	0.775	0.976	0.769	0.984	0.795	1	0.780	1
8	-5.194	0.820	0.764	0.964	0.759	0.973	0.783	1	0.769	1
9	-5.211	0.806	0.759	0.961	0.755	0.972	0.775	1	0.763	1
2	-5.218	0.801	0.740	0.934	0.735	0.942	0.760	1	0.745	1
7	-5.236	0.786	0.722	0.913	0.716	0.920	0.743	1	0.728	1
6	-5.237	0.786	0.720	0.910	0.715	0.917	0.742	1	0.726	1
4	-5.267	0.762	0.711	0.901	0.707	0.910	0.728	1	0.716	1

**Table 3**

90% Confidence Intervals for Technical Efficiency

Texas Utilities

IV Estimation Using Three Lagged Values

Firm No.	$\hat{\theta}_i$	TE <sub>i</sub>	MargCB		MCB		1-Sided MargCB		1-Sided MCB	
			LB	UB	LB	UB	LB	UB	LB	UB
5	-4.988	1	0.943	1	0.929	1	0.993	1	0.953	1
3	-5.078	0.915	0.789	1	0.780	1	0.830	1	0.803	1
10	-5.140	0.859	0.760	0.982	0.753	0.987	0.793	1	0.771	1
1	-5.170	0.834	0.765	0.984	0.761	0.992	0.788	1	0.773	1
8	-5.188	0.819	0.756	0.972	0.751	0.980	0.777	1	0.763	1
9	-5.204	0.806	0.752	0.972	0.749	0.982	0.770	1	0.759	1
2	-5.212	0.800	0.730	0.942	0.725	0.949	0.754	1	0.738	1
6	-5.231	0.785	0.709	0.915	0.704	0.921	0.734	1	0.718	1
7	-5.231	0.785	0.713	0.920	0.708	0.926	0.737	1	0.721	1
4	-5.261	0.761	0.703	0.908	0.699	0.916	0.723	1	0.710	1

**Table 4**

90% Confidence Intervals for Technical Efficiency  
 Indonesian Rice Farms  
 Robust Variance Matrix for the Effects

Firm No.	$\hat{\theta}_i$	TE <sub>i</sub>	MargCB		MCB		1-Sided MargCB		1-Sided MCB	
			LB	UB	LB	UB	LB	UB	LB	UB
164	5.556	1	0.660	1	0.487	1	0.781	1	0.522	1
118	5.486	0.932	0.573	1	0.424	1	0.678	1	0.452	1
163	5.484	0.930	0.574	1	0.425	1	0.678	1	0.453	1
:	:	:	:	:	:	:	:	:	:	:
31	5.072	0.616	0.375	1	0.275	1	0.445	1	0.294	1
:	:	:	:	:	:	:	:	:	:	:
15	4.966	0.554	0.339	1	0.250	1	0.402	1	0.267	1
:	:	:	:	:	:	:	:	:	:	:
16	4.859	0.498	0.303	1	0.223	1	0.360	1	0.238	1
:	:	:	:	:	:	:	:	:	:	:
117	4.586	0.379	0.231	1	0.170	1	0.274	1	0.181	1
45	4.550	0.365	0.224	1	0.165	1	0.265	1	0.176	1

## REFERENCES

- Arellano, M. (1990), "Testing for Autocorrelation in Dynamic Random Effects Models," *Review of Economic Studies* 57, 127-134.
- Dunnett, C.W. (1955), "A Multiple Comparison Procedure for Comparing Several Treatments with a Control," *Journal of the American Statistical Association* 50, 1096-1121.
- Dunnett, C.W. (1964), "New Tables for Multiple Comparisons with a Control," *Biometrics* 20, 482-491.
- Edwards, D.G. and J.C. Hsu (1983), "Multiple Comparisons with the Best Treatment," *Journal of the American Statistical Association* 78, 965-971. Corrigenda (1984), *Journal of the American Statistical Association* 79, 965.
- Erwidodo (1990), "Panel Data Analysis on Farm-Level Efficiency, Input Demand and Output Supply of Rice Farming in West Java, Indonesia," unpublished dissertation, Department of Agricultural Economics, Michigan State University.
- Gupta, S.S. (1965), "On Some Multiple Decision (Selection and Ranking) Rules," *Technometrics* 7, 225-245.
- Hochberg, Y. and A.C. Tamhane (1987), *Multiple Comparison Procedures*, New York: John Wiley and Sons.
- Horrace, W.C. (1998), "Tables of Percentages of the k-Variate Normal Distribution for Large Values of k," *Communications in Statistics: Simulation and Computation* 27, 823-831.
- Horrace, W.C. and P. Schmidt (1996), "Confidence Statements for Efficiency Estimates from Stochastic Frontier Models," *Journal of Productivity Analysis* 7, 257-282.
- Horrace, W.C. and P. Schmidt (2000), "Multiple Comparisons with the Best, with Economic Applications," *Journal of Applied Econometrics* 15, 1-26.
- Hsu, J.C. (1981), "Simultaneous Confidence Intervals for All Distances from the Best," *Annals of Statistics* 9, 1026-1034.
- Hsu, J.C. (1984), "Constrained Simultaneous Confidence Intervals for Multiple Comparisons with the Best," *Annals of Statistics* 12, 1145-1150.
- Hsu, J.C. (1996), *Multiple Comparisons: Theory and Methods*, London: Chapman and Hall.
- Koop, G., J. Osiewalski and M.F. Steel (1997), "Bayesian Efficiency Analysis through Individual Effects: Hospital Cost Frontiers," *Journal of Econometrics* 76, 77-106.

Kumbhakar, S.C. (1994), "Estimation of Cost Efficiency with Heteroscedasticity: An Application to Electric Utilities in Texas, 1966 - 1985," unpublished manuscript, Department of Economics, University of Texas at Austin.

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