ESTIMATES OF TECHNICAL INEFFICIENCY

IN STOCHASTIC FRONTIER MODELS WITH PANEL DATA:

GENERALIZED PANEL JACKKNIFE ESTIMATION

Panutat Satchachai
Michigan State University

Peter Schmidt
Michigan State University
Yonsei University

Estimates of technical inefficiency based on fixed effects estimation of the stochastic frontier model with panel data are biased upward. Previous work has attempted to correct this bias using the bootstrap, but in simulations the bootstrap corrects only part of the bias. The usual panel jackknife is based on the assumption that the bias is of order T^{-1} and is similar to the bootstrap. We show that when there is a tie or a near tie for the best firm, the bias is of order $T^{-1/2}$, not T^{-1}, and this calls for a different form of the jackknife. The generalized panel jackknife is quite successful in removing the bias. However, the resulting estimates have a large variance.

July 16, 2009

1. INTRODUCTION

In this paper we consider the stochastic frontier model with time-invariant technical inefficiency in a panel data setting. This model was first considered by Pitt and Lee (1981), who estimated the model by MLE given a distributional assumption for technical inefficiency. Without such a distributional assumption, Schmidt and Sickles (1984) proposed fixed effects estimation. In this approach, the frontier intercept is estimated as the maximum of the estimated firm-specific intercepts, and a firm's level of inefficiency is measured by the difference between the frontier intercept and the firm's intercept.

It is well understood that the "max" operation causes the estimated frontier intercept, and therefore the estimated inefficiency levels, to be biased upward. Schmidt and Sickles (1984), Park and Simar (1994) and Kim, Kim and Schmidt (2007) discuss this problem. Hall, Härdle and Simar (1995) show that the bootstrap is asymptotically (as $T \to \infty$ with N fixed) valid in this setting, provided that there is a unique best firm (no tie for the largest population intercept), and Kim, Kim and Schmidt (2007) use the bootstrap to construct a bias-corrected estimate of the frontier intercept (and therefore of inefficiency levels). The bootstrap is used to estimate the bias, which is then subtracted from the original estimate. In their simulations, Kim, Kim and Schmidt (2007) found that the bias correction was partially successful. It removed some but not all of the bias. Often it seemed to remove about half of the bias. Why it removed *half* of the bias, as opposed to some other fraction, is an interesting puzzle.

In this paper we consider bias corrections based on the jackknife. There are two motivations for doing so. Our first motivation is that the jackknife is thought to be similar to the bootstrap, but it is analytically simpler. Therefore we use the jackknife to explain why it is that under certain circumstances we remove half of the bias, and so we at least partially resolve the puzzle of the previous paragraph. The second motivation is to investigate whether the jackknife is practically useful as a bias-reduction technique in this model. Here we are less successful, because the jackknife does effectively remove the bias of the estimate, but the variance and MSE of the jackknife estimate are unfortunately rather large.

The appropriate form the jackknife depends on the order of the leading term in an expansion of the bias of the estimate. If the bias of the estimate is of order T^{-1}, the usual delete-one panel jackknife estimator (as in Hahn and Newey (2004)) should remove the bias. However, intuitively we would expect the jackknife bias correction to be similar to the bootstrap bias correction, which was only partially successful. Thus it would seem that the finite-sample relevance of the bias being of order T^{-1} may be questionable.

In this paper we analyze the case of an exact tie for the best firm. In this case the bootstrap is not asymptotically valid. Furthermore, we show that the bias of the fixed effects estimate of the frontier intercept is of order $T^{-1/2}$, not T^{-1}. In this case the usual delete-one panel jackknife does not properly remove the bias. Indeed, we show that it removes (approximately) half of the bias. A different form of the jackknife, which we call the generalized panel jackknife, does remove the bias in this case.

In the simulations of Kim, Kim and Schmidt (2007) there was not an exact tie, and an exact tie may also be unlikely in actual data. However, if there is nearly a tie, in the sense that there is substantial uncertainty *ex post* about which is the best firm, it is not clear whether asymptotics that assume no tie are more relevant than asymptotics that assume an exact tie. In order to further analyze a near tie, we give a specific definition (involving a local parameterization) of a “near tie,” and we show that the bias is again of order $T^{-1/2}$, so that the generalized panel jackknife is needed to successfully remove the bias.

We then perform simulations to assess the finite-sample relevance of these results.

The plan of the paper is as follows. In Section 2, we define some notation and give a brief review of fixed effects estimation of the stochastic frontier model with panel data. In Section 3 we show that the bias is of order $T^{-1/2}$ for the case of an exact tie or a "near tie." Section 4 describes the generalized panel jackknife that is appropriate in this circumstance. In Section 5 we explain the design of our Monte Carlo experiments, and Section 6 gives its results. Finally, Section 7 contains our concluding remarks.

2. FIXED EFFECTS ESTIMATION OF THE MODEL

Consider a single-output production function with time-invariant technical inefficiency $u_i \geq 0$. There are N firms, indexed by $i = 1,...,N$, over T time periods, indexed by $t = 1,...,T$. We consider the linear regression model of Schmidt and Sickles (1984):

$$y_{it} = \alpha + x_{it}'\beta + v_{it} - u_i, i = 1,...,N; t = 1,...,T, \tag{1}$$

where y_{it} is the logarithm of output for firm i at time t; x_{it} is a vector of K inputs (e.g., in logarithms for a Cobb-Douglas production function); β is a $K \times 1$ vector of coefficients; and v_{it} is an i.i.d. idiosyncratic error with mean zero and finite variance. The v_{it} represent uncontrollable shocks that affect the level of output, e.g., luck, weather, or machine performance. The time-invariant technical inefficiency u_i satisfies $u_i \geq 0$ for all i and $u_i > 0$ for some i. There is no distributional assumption on u_i except that it is one-sided.

Defining $\alpha_i = \alpha - u_i$, we can write (1) as a standard panel data model:

$$y_{it} = \alpha_i + x_{it}'\beta + v_{it}. \tag{2}$$

Obviously, $\alpha_i \leq \alpha$ since $u_i \geq 0$. When α_i (and u_i) is treating as fixed, (2) leads to a fixed effects estimation problem in which neither a distribution for technical inefficiency nor the independence between technical inefficiency and x_{it} or v_{it} (or both) is needed. We assume strict exogeneity of the regressors x_{it} in the sense that $(x_{i1},...,x_{iT})'$ is independent of $(v_{i1},...,v_{iT})'$. There is no restriction on the distribution of v_{it} other than zero mean and finite variance.

To estimate β, we use the fixed effects estimate $\hat{\beta}$, which can be estimated as "least squares with dummy variables," by regressing y_{it} on x_{it} and a set of N dummy variables, or as the "within estimator," by regressing $(y_{it} - \bar{y}_i)$ on $(x_{it} - \bar{x}_i)$. Given the estimate $\hat{\beta}$, the estimates $\hat{\alpha}_i$ can be recovered as the averages of the firm-specific residuals, i.e., $\hat{\alpha}_i = \bar{y}_i - \bar{x}_i'\hat{\beta}$ where $\bar{y}_i = T^{-1}\sum_t y_{it}$ and $\bar{x}_i = T^{-1}\sum_t x_{it}$, or equivalently as the coefficients of the firm-specific dummy variables.

The within estimator $\hat{\beta}$ is consistent as N or $T \to \infty$, and the firm-specific intercepts $\hat{\alpha}_i$ are consistent as $T \to \infty$. To estimate α and u_i, Schmidt and Sickles (1984) suggested the following estimators:

$$\hat{\alpha} = \max_{j=1,...,N} \hat{\alpha}_j, \qquad \hat{u}_i = \hat{\alpha} - \hat{\alpha}_i, \qquad i = 1,...,N. \tag{3}$$

Park and Simar (1994) show that these estimates are consistent as $N \to \infty$, $T \to \infty$, and $T^{-1/2}\ln(N) \to 0$.

In this paper, to maintain the connection to the earlier literature on bootstrapping of this model, and also the literature on the jackknife, we will consider asymptotic arguments as $T \to \infty$ with N fixed. In this case we can only measure inefficiency relative to the best of the N firms.

For ease of presentation, we follow Kim, Kim and Schmidt (2007) and rank the intercepts α_i such that $\alpha_{(1)} \le \alpha_{(2)} \le ... \le \alpha_{(N)}$, so that (N) indexes the firm with the largest value of α_i among the N firms, which we will call the best firm. Similarly, we rank the levels of technical inefficiency u_i in the opposite order such that $u_{(1)} \ge u_{(2)} \ge ... \ge u_{(N)}$. Obviously, $\alpha_{(i)} = \alpha - u_{(i)}$ for all i and specifically $\alpha_{(N)} = \alpha - u_{(N)}$.

Now we define the relative inefficiency measures

$$u_i^* = u_i - u_{(N)} = \alpha_{(N)} - \alpha_i . \quad (4)$$

These are the focus of this paper since, as $T \to \infty$ with N fixed, $\hat{\alpha}$ is a consistent estimate of $\alpha_{(N)}$, not α, and $\hat{u}_i^*$ is a consistent estimate of u_i^*, not u_i.

Although $\hat{\alpha}$ is consistent for $\alpha_{(N)}$ (as $T \to \infty$ with N fixed), it is biased upward for finite T. This is true because $\hat{\alpha} \ge \hat{\alpha}_{(N)}$ and $E(\hat{\alpha}_{(N)}) = \alpha_{(N)}$. That is, the max operator in (3) induces an upward bias: the largest $\hat{\alpha}_i$ is more likely to contain positive estimation error than negative error. The upward bias in the estimate $\hat{\alpha}$ induces an upward bias in the estimates of relative technical inefficiency. That is, $E(\hat{\alpha}) - \alpha_{(N)} = E(\hat{u}_i) - u_i^*$. Therefore we will simply evaluate the bias of $\hat{\alpha}$ as an estimate of $\alpha_{(N)}$; there is no need to separately evaluate the bias of the estimates of relative technical inefficiency.

The bias of $\hat{\alpha}$ as an estimate of $\alpha_{(N)}$ corresponds to what Kim, Kim and Schmidt (2007) call the "first-level bias." To correct this first-level bias, Kim, Kim and Schmidt (2007) consider a bootstrap bias correction for the fixed effects estimate. They evaluate the "second-level bias," $E(\hat{\alpha}^{boot}) - \hat{\alpha}$, and use it to correct the first-level bias. That is, if the second-level bias equals the first-level bias, we would want to evaluate

$$\hat{\alpha} - [E(\hat{\alpha}^{boot}) - \hat{\alpha}] = 2\hat{\alpha} - E(\hat{\alpha}^{boot}) . \tag{5}$$

The feasible version of this is

$$\hat{\alpha}^{boot}_{BC} = 2\hat{\alpha} - B^{-1}\sum_{b=1}^{B} \hat{\alpha}^{(b)}, \tag{6}$$

where "b" represents a single bootstrap replication and "B" is the total number of bootstrap replications. In their simulations (see their Table 4), this estimate removes some but not all of the bias in $\hat{\alpha}$. Often it seems to remove about half of the bias. As noted in the introduction, the fact that *half* of the bias is removed is the puzzle that at least partially motivated our interest in the jackknife.

3. DERIVING THE ORDER IN PROBABILITY OF THE BIAS

In this section, we show that the bias of $\hat{\alpha}$ is of order T^{-1} if there is no tie for the best firm; that is, if $\alpha_{(N)}$ is strictly larger than all of the other α_i. However, if there is a tie for the best firm, or if there is a "near tie" (in a sense defined precisely below), the bias is of order $T^{-1/2}$.

For simplicity, we will discuss the simple case of no regressors:

$$y_{it} = \alpha_i + v_{it}, i = 1,\ldots,N; t = 1,\ldots,T, \tag{7}$$

where v_{it} are i.i.d. with mean zero and variance σ^2. Thus $\hat{\alpha}_i = \bar{y}_i$. The various $\hat{\alpha}_i$ are independent and $\sqrt{T}(\hat{\alpha}_i - \alpha_i) \rightarrow N(0,\sigma^2)$. However, the inclusion of regressors would not alter our results since the within estimator of β is unbiased, and our results really only require that the vector whose i^{th} element is $\sqrt{T}(\hat{\alpha}_i - \alpha_i)$ is normal with mean zero and finite variance matrix. See Hall, Härdle and Simar (1995), Appendix (i), equation (A.1) for this condition, which would still hold with regressors.

3.1 The Case of No Tie

Suppose first that there is no tie for the best firm. That is, there is a unique firm "i" such that $\alpha_{(N)} = \alpha_i$.

Hall, Härdle and Simar (1995) show the equivalence of (i) there is no tie for the best firm, and (ii) the asymptotic distribution of $\hat{\alpha}$ is normal. More precisely, they show that if there is no tie, $P(\hat{\alpha} = \hat{\alpha}_{(N)}) \to 1$ as $T \to \infty$, so that the asymptotic distribution of $\hat{\alpha}$ is the same as the asymptotic distribution of $\hat{\alpha}_{(N)}$, the estimate of $\alpha_{(N)}$ that would be used if the identity of the best firm were known. Since $\hat{\alpha}_{(N)}$ is unbiased, it follows that $\sqrt{T}$ times the bias of $\hat{\alpha}$ must go to zero as $T \to \infty$. Thus we conclude that the bias of $\hat{\alpha}$ is of an order smaller than $T^{-1/2}$. We presume that it is of order T^{-1}.

3.2 The Case of an Exact Tie

Suppose now that there is a tie for the best firm (the largest α_i). Specifically suppose that the first "k" firms are tied, so that $\alpha_{(N)} = \alpha_1 = \alpha_2 = ... = \alpha_k$ for $2 \le k \le N$. Again the discussion in Hall, Härdle and Simar (1995, Appendix (i)) applies. With a probability that approaches one as $T \to \infty$, $\hat{\alpha}$ will equal $\hat{\alpha}_i$ for some i with $1 \le i \le k$, that is, the estimated best firm will be one of the k truly best firms. Therefore with a probability that approaches one,

$$\begin{aligned}\sqrt{T}(\hat{\alpha} - \alpha_{(N)}) &= \sqrt{T}\max(\hat{\alpha}_1 - \alpha_{(N)}, \hat{\alpha}_2 - \alpha_{(N)}, ..., \hat{\alpha}_k - \alpha_{(N)}) \\ &= \max\{\sqrt{T}(\hat{\alpha}_1 - \alpha_{(N)}), \sqrt{T}(\hat{\alpha}_2 - \alpha_{(N)}), ..., \sqrt{T}(\hat{\alpha}_k - \alpha_{(N)})\}\end{aligned} \quad (8)$$

and therefore $\sqrt{T}(\hat{\alpha} - \alpha_{(N)}) \rightarrow Z$ where Z is the maximum of a set of k normals with zero mean. For $k > 1$, Z is not normal, and $E(Z) > 0$. The bias of $\hat{\alpha}$ is therefore, for large T, $T^{-1/2}E(Z)$, which is of order $T^{-1/2}$.

We can give an explicit expression for the case of $N = k = 2$ and the simple model above (with no regressors). We first state the following Lemma.

Lemma 1 Suppose X_1 and X_2 are i.i.d. $N(\mu, \sigma^2)$, i.e.,

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left[\begin{pmatrix} \mu \\ \mu \end{pmatrix}, \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}\right],$$

then

$$E[\max(X_1, X_2) - \mu] = (1/\sqrt{\pi})\sigma. \tag{9}$$

Proof. Let

$$\begin{pmatrix} Y \\ Z \end{pmatrix} = \begin{pmatrix} X_1 \\ X_1 - X_2 \end{pmatrix} \sim N\left[\begin{pmatrix} \mu \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 & \sigma^2 \\ \sigma^2 & 2\sigma^2 \end{pmatrix}\right].$$

So, $\rho = \sigma^2 / \sqrt{2\sigma^2\sigma^2} = 1/\sqrt{2}$ and

$$E(X_1 \mid X_1 > X_2) = E(Y \mid Z > 0)$$

$$= \mu + (1/\sqrt{2})\sigma\lambda(0), \text{ where } \lambda(\cdot) \text{ is the normal hazard function}$$

$$= \mu + (1/\sqrt{2})\sigma(\sqrt{2/\pi}), \text{ since } \lambda(0) = \phi(0)/(1 - \Phi(0)) = \sqrt{2/\pi}$$

$$= \mu + (1/\sqrt{\pi})\sigma.$$

Hence, $E(X_1 \mid X_1 > X_2) - \mu = (1/\sqrt{\pi})\sigma$ and

$$E[\max(X_1, X_2)] = (1/2)E(X_1 \mid X_1 > X_2) + (1/2)E(X_2 \mid X_2 > X_1), \text{ by symmetry}$$

$$= E(X_1 \mid X_1 > X_2), \text{ since } X_1 \text{ and } X_2 \text{ are i.i.d.}$$

Therefore, $bias = E[\max(X_1, X_2)] - \mu = (1/\sqrt{\pi})\sigma$. □

In the present setting, "X_1" and "X_2" are $\hat{\alpha}_1$ and $\hat{\alpha}_2$, "μ" $= \alpha_1 = \alpha_2$, the variance "σ^2" is σ^2/T, and the bias of $\hat{\alpha} = \max(\hat{\alpha}_1, \hat{\alpha}_2)$ equals $(1/\sqrt{\pi})T^{-1/2}\sigma$. Clearly, this is proportional to $T^{-1/2}$.

3.3 The Case of a Near Tie

In the previous sections we saw that the bias of $\hat{\alpha}$ is of order T^{-1} if there is no tie for the best firm, while it is of order $T^{-1/2}$ if there is an exact tie. It is not clear how relevant either set of results will be in finite samples if there is (in some sense) nearly a tie. Intuitively that will depend on how close we are to a tie, which depends not only on how close the α_i are to each other, but also on $T^{-1/2}\sigma$, which is the standard deviation of the $\hat{\alpha}_i$.

One way to model this is by a "local to tie" parameterization. So, to keep things simple, let $N = 2$, $\alpha_1 > \alpha_2$, and $\alpha_2 = \alpha_1 - T^{-1/2}c$ for $c > 0$, where c does not depend on T. Then in our simple (no regressors) model, $\sqrt{T}(\hat{\alpha}_1 - \alpha_1) \to N(0, \sigma^2)$. Also $\sqrt{T}(\hat{\alpha}_2 - \alpha_2) \to N(0, \sigma^2)$ and so $\sqrt{T}(\hat{\alpha}_2 - \alpha_1 + T^{-1/2}c) \to N(0, \sigma^2)$, or $\sqrt{T}(\hat{\alpha}_2 - \alpha_1) \to N(-c, \sigma^2)$. Then

$$\begin{aligned} \sqrt{T}[\max(\hat{\alpha}_1, \hat{\alpha}_2) - \alpha_1] &= \max[\sqrt{T}(\hat{\alpha}_1 - \alpha_1), \sqrt{T}(\hat{\alpha}_2 - \alpha_1)] \\ &\to Z \end{aligned} \tag{10}$$

where "Z" is the max of a $N(0,\sigma^2)$ random variable and a $N(-c,\sigma^2)$ random variable. Clearly $E(Z) \geq E(N(0,\sigma^2)) = 0$ and the bias of $\hat{\alpha}$ is again (for large T) $T^{-1/2}E(Z)$, which is of order $T^{-1/2}$.

A similar analysis applies if $\alpha_2 = \alpha_1 - T^{-\gamma}c$ where $c > 0$ and $\gamma \geq 1/2$. The value of c matters (as above) when $\gamma = 1/2$ but it does not affect the limit distribution if $\gamma > 1/2$. So the asymptotics for the case of a "near tie" are very similar to those for an exact tie if a tie is near enough.

Once again we can give an explicit expression for the case of $N = k = 2$ and the simple model (no regressors). We state without proof the following Lemma.

Lemma 2 Let X_1 and X_2 be independent normals, where $X_1 \sim N(0,\sigma^2)$ and $X_2 \sim N(\mu_2,\sigma^2)$. Then

$$E[\max(X_1, X_2)] = [\Phi(\mu_*/\sqrt{2})\mu_* + \sqrt{2}\phi(\mu_*/\sqrt{2})]\sigma, \tag{11}$$

where $\mu_* = \mu_2/\sigma$.

To apply this to our model, "X_1" and "X_2" are $\hat{\alpha}_1$ and $\hat{\alpha}_2$, "σ^2" is σ^2/T, $\mu_2 = -T^{-1/2}c$, and $\mu_* = -T^{-1/2}c/T^{-1/2}\sigma = -c/\sigma$. So the bias is

$$bias = [\Phi(-c/\sqrt{2}\sigma)(-c/\sigma) + \sqrt{2}\phi(-c/\sqrt{2}\sigma)]T^{-1/2}\sigma, \tag{12}$$

which is indeed proportional to $T^{-1/2}$.

4. CORRECTING BIAS WITH THE PANEL JACKKNIFE AND THE GENERALIZED PANEL JACKKNIFE

4.1 The Panel Jackknife

Jackknife estimation is an automatic bias reduction tool under the assumption of the existence of a series expansion for the bias of an estimator. Quenouille (1956) and Tukey (1958) show that using the jackknife estimates based on removing data and then recalculating the estimator removes the first order bias from an initial estimator. For a basic background discussion of jackknife estimation, see Miller (1974).

To describe the jackknife in a general setting, let the data be indexed by $t = 1, 2, ..., T$. Let $\hat{\theta}$ be the estimator based on all T observations, and let $\hat{\theta}_{(t)}$ be the "delete-observation-*t*" estimator that omits observation t and uses the other $T-1$ observations. Then the jackknife estimator is

$$J(\hat{\theta}) = T\hat{\theta} - (T-1)T^{-1}\sum_t \hat{\theta}_{(t)}. \tag{13}$$

This estimator is said to remove the bias of order T^{-1}, in the following sense. Suppose that

$$E(\hat{\theta}) = \theta + T^{-1}B + T^{-2}D + O(T^{-3}) . \tag{14}$$

Then

$$E[J(\hat{\theta})] = \theta + \left(\frac{1}{T} - \frac{1}{T-1}\right)D + O(T^{-2}) = \theta + O(T^{-2}) . \tag{15}$$

So if the bias is of order T^{-1}, in the sense that (14) holds, the jackknife leaves only the bias of order T^{-2}.

Hahn and Kuersteiner (2004), Hahn and Newey (2004), and Fernández-Val and Vella (2007) apply the jackknife to nonlinear panel data models and dynamic panel data models. In the panel data setting, even though there are really NT observations, we treat the number of observations in (13) as T, and to calculate $\hat{\theta}_{(t)}$ we delete the t^{th} period observation for each cross-sectional unit. (This is done because, in the models they consider, the bias is of order T^{-1}.) We refer to this procedure as the "panel jackknife."

Other similar versions of the jackknife can remove bias of order T^{-1}. For example, Dhaene, Jochmans and Thuysbaert (2006) propose the half-panel jackknife estimator:

$$J(\hat{\theta}^{half-panel}) = 2\hat{\theta} - (1/2)(\hat{\theta}_1 + \hat{\theta}_2)\,, \tag{16}$$

where $\hat{\theta}$ is the fixed effects estimator based on the full sample; $\hat{\theta}_1$ and $\hat{\theta}_2$ are based on the first- and second- halves of the panel sample, where each half-panel consists of $T/2$ consecutive observations over time for all cross-sectional units. They show that the half-panel jackknife estimator also removes the bias of order T^{-1} from the original estimator. However, for the case that the bias is of order T^{-1}, we will consider only the standard panel jackknife as described above.

It is obvious that when there is no tie, the panel jackknife will remove the first-level bias of the estimate of $\alpha_{(N)}$ (hence, the bias of the estimates of relative technical inefficiency u_i^*) since the bias is of order T^{-1}. For the cases of an exact tie and a near tie, however, we need a jackknife estimator that can handle bias of order $T^{-1/2}$. The difference in the order of the bias leads us to the generalized jackknife.

4.2 The Generalized Jackknife

Schucany, Gray and Owen (1971) were the first to propose a jackknife estimator that can handle a more general form of bias. It was not until later that Gray and Schucany (1972) gave it the name "generalized jackknife." Gray and Schucany (1972) define the generalized jackknife as the following.

Definition Gray and Schucany (1972)'s Definition 2.1. Let $\hat{\theta}_1$ and $\hat{\theta}_2$ be two estimators for θ. Then, for any real number $R \neq 1$, the generalized jackknife estimator $G(\hat{\theta}_1, \hat{\theta}_2)$ is defined as

$$G(\hat{\theta}_1, \hat{\theta}_2) = \frac{\hat{\theta}_1 - R\hat{\theta}_2}{1 - R}. \tag{17}$$

The usual (Quenouille) jackknife corresponds to $\hat{\theta}_1 = \hat{\theta}$, $\hat{\theta}_2 = T^{-1}\sum_t \hat{\theta}_{(t)}$, and $R = (T-1)/T$.

If we can express the bias of the estimators in terms of the sample size T and the true parameter θ, we can choose R so that the generalized jackknife is unbiased.

Theorem 1 Gray and Schucany (1972)'s Theorem 2.1. If the bias of the estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ can be expressed as

$$E(\hat{\theta}_k) = \theta + b_k(T, \theta), k = 1, 2;$$

$$b_2(T, \theta) \neq 0;$$

and

$$R = \frac{b_1(T, \theta)}{b_2(T, \theta)} \neq 1,$$

then

$$E[G(\hat{\theta}_1, \hat{\theta}_2)] = \theta.$$

Proof.

$$E[G(\hat{\theta}_1,\hat{\theta}_2)] = \frac{[\theta + b_1(T,\theta)] - R[\theta + b_2(T,\theta)]}{1-R}$$

$$= \theta + \frac{b_1(T,\theta) - Rb_2(T,\theta)}{1-R}$$

$$= \theta, \text{since } Rb_2(T,\theta) = b_1(T,\theta).$$

□

In general, we do not have a bias expression of the form of the previous theorem, but we have a series expansion of the bias with a leading term of known order. Then the generalized jackknife removes the leading term of the series expansion of the bias.

Theorem 2 Gray and Schucany (1972)'s Theorem 2.2. If the bias of the estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ can be expanded as an infinite series:

$$E(\hat{\theta}_k) = \theta + \sum_{i=1}^{\infty} b_{ki}(T,\theta), k = 1,2$$

and

$$R = \frac{b_{11}(T,\theta)}{b_{21}(T,\theta)} \neq 1,$$

then

$$E[G(\hat{\theta}_1,\hat{\theta}_2)] = \theta + \frac{\sum_{i=2}^{\infty} b_{1i}(T,\theta) - R\sum_{i=2}^{\infty} b_{2i}(T,\theta)}{1-R}.$$

Proof. Similar to the proof of Theorem 1. □

4.3 The Generalized Panel Jackknife When the Bias Is of Order $T^{-1/2}$

We are specifically interested in the case that the bias of $\hat{\theta}$ is of order $T^{-1/2}$. Suppose that the following expansion holds:

$$E(\hat{\theta}) = \theta + T^{-1/2}B + T^{-1}D + O(T^{-3/2}). \tag{18}$$

As before, we let $\hat{\theta}_1 = \hat{\theta}$ (T observations) and $\hat{\theta}_2 = T^{-1}\sum_t \hat{\theta}_{(t)}$. Then the weight R in Theorem 2 is equal to

$$R = (B/\sqrt{T})/(B/\sqrt{T-1}) = \sqrt{T-1}/\sqrt{T} \tag{19}$$

and the generalized jackknife is

$$G(\hat{\theta}) = \frac{\sqrt{T}}{\sqrt{T}-\sqrt{T-1}}\hat{\theta} - \frac{\sqrt{T-1}}{\sqrt{T}-\sqrt{T-1}}T^{-1}\sum_t \hat{\theta}_{(t)} . \tag{20}$$

It is then easy to verify that the bias of $G(\hat{\theta})$ is of order T^{-1}; that is, the $T^{-1/2}$ term in the bias of $\hat{\theta}$ has been removed.

In the panel data case, once again we treat the number of observations as T, and $\hat{\theta}_{(t)}$ is calculated by deleting the t^{th} time period observation for each cross-sectional unit. We will call this the generalized panel jackknife.

The generalized jackknife removes bias more aggressively than the usual jackknife, in the sense that the weights attached to $\hat{\theta}$ and to $T^{-1}\sum_t \hat{\theta}_{(t)}$ are larger. For example, for $T = 10$ we have

$$J(\hat{\theta}) = 10\hat{\theta} - 9(T^{-1}\sum_t \hat{\theta}_{(t)})$$

$$G(\hat{\theta}) = 19.5\hat{\theta} - 18.5(T^{-1}\sum_t \hat{\theta}_{(t)}).$$

Similarly for $T = 50$ we have

$$J(\hat{\theta}) = 50\hat{\theta} - 49(T^{-1}\sum_t \hat{\theta}_{(t)})$$

$$G(\hat{\theta}) = 99.5\hat{\theta} - 98.5(T^{-1}\sum_t \hat{\theta}_{(t)}).$$

A detail that we do not pursue in this paper is that it is possible to consider a second level of the jackknife. If the bias of the original estimate is of order $T^{-1/2}$, the bias of $G(\hat{\theta})$ is of order T^{-1}. The usual panel jackknife applied to the estimator $G(\hat{\theta})$ would remove the bias of order T^{-1}. The resulting estimator would be a linear combination of the original estimate, the various "drop one observation" estimates, and the various "drop two observations" estimates.

4.4 What If The Wrong Jackknife Is Used?

We have seen that the usual panel jackknife is appropriate when the bias is of order T^{-1}, whereas the generalized panel jackknife is appropriate when the bias is of order $T^{-1/2}$. This raises the question of what happens if the wrong version of the jackknife is used.

Theorem 3 If the bias of $\hat{\theta}$ is of order $T^{-1/2}$, the usual panel jackknife corrects approximately half of the bias.

Proof. We have $E(\hat{\theta}) = \theta + T^{-1/2}B + \textit{higher order terms}$. So, dropping the higher order terms, we calculate

$$E[J(\hat{\theta})] = \theta + B\sqrt{T} - B\sqrt{T-1} = \theta + \frac{B}{\sqrt{T} + \sqrt{T-1}}. \qquad (21)$$

Comparing the bias in this expression to the original bias of $T^{-1/2}B$, we have removed about half of the first-order bias term. □

Theorem 3 is our explanation of the puzzle that the bootstrap and the jackknife often correct half of the bias. Cases where this occurs correspond to more or less an exact tie.

Theorem 4 If the bias of $\hat{\theta}$ is of order T^{-1}, the bias of the generalized panel jackknife is approximately the negative of the bias of the original estimate.

Proof. Suppose $E(\hat{\theta}) = \theta + T^{-1}B + \textit{higher order terms.}$ So, again dropping the higher order terms,

$$E[G(\hat{\theta})] = \frac{\sqrt{T}}{\sqrt{T}-\sqrt{T-1}}[\theta + T^{-1}B] - \frac{\sqrt{T-1}}{\sqrt{T}-\sqrt{T-1}}T^{-1}\sum_t [\theta + (T-1)^{-1}B]$$

$$= \frac{1}{\sqrt{T}-\sqrt{T-1}}[\theta\sqrt{T} + T^{-1/2}B - \theta\sqrt{T-1} - (T-1)^{-1/2}B]$$

$$= \theta + \frac{1}{\sqrt{T}-\sqrt{T-1}}[T^{-1/2} - (T-1)^{-1/2}]B \tag{22}$$

$$= \theta + \frac{1}{\sqrt{T}-\sqrt{T-1}}\left(\frac{\sqrt{T-1}-\sqrt{T}}{\sqrt{T}\sqrt{T-1}}\right)B$$

$$= \theta - (\sqrt{T}/\sqrt{T-1})T^{-1}B.$$

So the bias of $G(\hat{\theta})$, $-(\sqrt{T}/\sqrt{T-1})T^{-1}B$, is approximately the negative of the original bias, $T^{-1}B$. □

5. DESIGN OF THE MONTE CARLO EXPERIMENTS

In this section, we conduct Monte Carlo simulations to investigate the finite sample performance of the following estimators of $\alpha_{(N)}$: (i) $\hat{\alpha}$, the maximum of the fixed effects estimates; (ii) $J(\hat{\alpha})$, the panel jackknife estimate; (iii) $G(\hat{\alpha})$, the generalized panel jackknife estimate; and (iv) $\hat{\alpha}_{BC}^{boot}$, the bias-corrected bootstrap point estimate.

We are primarily interested in the bias of these estimators. However, we will also report their variance and mean square error. These measures are defined precisely later in this section.

The model is the simple panel data model with no regressors, as given in (7). Thus, the data generating process is

$$\begin{aligned} y_{it} &= \alpha + v_{it} - u_i \\ &= \alpha_i + v_{it},\ i = 1,\dots,N;\ t = 1,\dots,T, \end{aligned} \tag{23}$$

where $\alpha_i = \alpha - u_i$; the u_i are i.i.d. half-normal: $u_i = |U_i|$ where $U_i \sim N(0,\sigma_u^2)$; and the v_{it} are normal with mean zero and variance σ_v^2. These distributional assumptions are not used in estimation. They just characterize the process that generates the data.

The set of parameters is $\{\alpha, \sigma_v^2, \sigma_u^2, N, T\}$ but this can be reduced somewhat. All of the results (bias, variance, and MSE) are invariant with respect to α, so we set it equal to one, without loss of generality. Also, only ratios of variances matter. If we multiply both σ_u^2 and σ_v^2 by a constant q, the biases of the estimates change by $\sqrt{q}$ and the MSE's change by q. So we really only need to consider three parameters: $N, T,$ and a relative variance parameter. Kim, Kim and Schmidt (2007) used the relative variance parameter $\gamma^* = (\sigma_u^2)_* / [\sigma_v^2 + (\sigma_u^2)_*]$, where $(\sigma_u^2)_* = \mathrm{var}(u) = ((\pi-2)/\pi)\sigma_u^2$. We will use instead the parameter μ_* defined by

$$\mu_* = \frac{(\sigma_u)_*}{T^{-1/2}\sigma_v}. \tag{24}$$

This is not a matter of substance. We use μ_* because we find it easier to interpret. It measures the standard deviation of the α_i in units of the standard deviation of the $\hat{\alpha}_i$. Also, for reasons given below, with this parameterization it turns out that T does not matter very much. Only μ_* and N turn out to be important.

So, in the end, our parameter space is $\{\mu_*, N, T\}$. We set scale by setting $\sigma_v^2 / T = 0.1$, which for a given T determines σ_v^2. Then, for a given μ_*, $(\sigma_u^2)_*$ is determined. We consider $\mu_* = 10^{-1}, 10^{-1/2}, 1, 10^{1/2}$, and 10. With $\sigma_v^2 / T = 0.1$, for a given value of μ_*, the values of $(\sigma_u^2)_*$ and σ_u^2 are as follows:

(1) $\mu_* = 10^{-1} = 0.1$: $(\sigma_u^2)_* = 0.001$; $\sigma_u^2 = 0.0028$;

(2) $\mu_* = 10^{-1/2} = 0.3162$: $(\sigma_u^2)_* = 0.01$; $\sigma_u^2 = 0.0275$;

(3) $\mu_* = 1$: $(\sigma_u^2)_* = 0.1$; $\sigma_u^2 = 0.2752$;

(4) $\mu_* = 10^{1/2} = 3.1623$; $(\sigma_u^2)_* = 1$; $\sigma_u^2 = 2.7519$;

(5) $\mu_* = 10$: $(\sigma_u^2)_* = 10$; $\sigma_u^2 = 27.5194$.

We consider sample sizes $N = 2, 10, 20, 50,$ and 100, and we set $T = 10$. We also considered $T = 5, 20, 50,$ and 100, and the results for these values of T are available in a supplementary set of tables, available from the authors on request.

The basic outcomes that we would expect in the simulations are as follows. First, bias will be larger when N is larger, but the effect of N on the relative performance of the various bias-corrected methods is not obvious. Second, bias will be larger when μ_* is smaller, since then the variability of the α_i is smaller relative to the sampling variability of the $\hat{\alpha}_i$. We might expect the ordinary panel jackknife or the bootstrap to be better than the generalized jackknife when μ_*

is large (we are farther from a tie), and vice-versa. Third, conditional on μ_*, we do not expect T to be very important. When we change T in our experiment, holding constant μ_* and σ_v^2/T, it means that σ_v^2 increases proportionally to T, and $(\sigma_u^2)_*$ is unchanged. Therefore neither the variability of the α_i nor the sampling variability of the $\hat{\alpha}_i$ changes. The only reason that T should matter is that the jackknife's weights on $\hat{\theta}$ and $T^{-1}\sum_t \hat{\theta}_{(t)}$ depend on T.

We consider three different variations of the setup we have just described.

Experiment I (No Tie). The setup of this experiment is exactly as just described. There are no restrictions on the α_i. They just follow from the draws of the half-normal u_i. This setup is very similar to that of Kim, Kim and Schmidt (2007).

Experiment II (Exact Tie). We generate data as described above. Now we (the data generator) know which firm is the best and the value $\alpha_{(N)}$ of its intercept. We randomly select one of the other $(N-1)$ firms and set its intercept also equal to $\alpha_{(N)}$. Therefore we have created an exact two-way tie for the best firm.

Experiment III (Near Tie). We start as in Experiment II. However, once we have observed the best firm and $\alpha_{(N)}$, we randomly select one of the other $(N-1)$ firms and set its intercept equal to

$$\alpha_{(N)} - T^{-1/2}[\alpha_{(N)} - \alpha_{(N-1)}]. \tag{25}$$

So, for example, if $T = 10$, we have now created a new second-best firm whose intercept is $\sqrt{10} = 3.162$ times closer to $\alpha_{(N)}$ than the previously second-best firm's intercept.

For each configuration of $\{\mu_*, N, T\}$, we perform 1,000 replications. Within each replication, the bias-corrected bootstrap estimate is based on 1,000 bootstrap replications.

For each of the estimators $\left(\hat{\alpha}, J(\hat{\alpha}), G(\hat{\alpha}), \hat{\alpha}_{BC}^{boot}\right)$ we calculate bias, variance and mean square error. The parameter being estimated, $\alpha_{(N)}$, varies across replications because of the random draws of the half-normal u_i that determine $\alpha_i = \alpha - u_i$. Therefore we will explicitly state our definition of bias, variance and MSE. First define: (i) $NREP =$ number of replications; (ii) $r =$ index of replication, $r = 1,\ldots,NREP$; (iii) $\theta_r =$ value of $\alpha_{(N)}$ in replication r; (iv) $\hat{\theta}_r =$ value of $\hat{\theta}$ in replication r (for any of the four estimators listed above); and (v) $\bar{\theta} = NREP^{-1}\sum_r \theta_r$ and $\bar{\hat{\theta}} = NREP^{-1}\sum_r \hat{\theta}_r$.

The definition of bias is straightforward:

$$bias(\hat{\theta}) = NREP^{-1}\sum_r (\hat{\theta}_r - \theta_r) = \bar{\hat{\theta}} - \bar{\theta}\,. \tag{26}$$

Then we define the mean squared error as

$$\begin{aligned} MSE(\hat{\theta}) &= NREP^{-1}\sum_r (\hat{\theta}_r - \theta_r)^2 \\ &= NREP^{-1}\sum_r [(\hat{\theta}_r - \theta_r) - bias(\hat{\theta})]^2 + bias(\hat{\theta})^2 \end{aligned} \tag{27}$$

and the variance as

$$\begin{aligned} \text{var}(\hat{\theta}) &= MSE(\hat{\theta}) - bias(\hat{\theta})^2 \\ &= NREP^{-1}\sum_r [(\hat{\theta}_r - \theta_r) - bias(\hat{\theta})]^2. \end{aligned} \tag{28}$$

6. RESULTS OF THE MONTE CARLO EXPERIMENTS

Tables 1, 2, and 3 give the results of Experiment I in which there is no tie. All of these results are for $T = 10$. Table 1 gives the bias of the estimates, while Table 2 gives variance and Table 3 gives MSE. In all three tables, column (1) gives results for $\hat{\alpha}$; column (2) gives results for the panel jackknife $J(\hat{\alpha})$; column (3) gives results for the generalized panel jackknife $G(\hat{\alpha})$; and column (4) gives results for the bias-corrected bootstrap point estimate $\hat{\alpha}_{BC}^{boot}$.

Consider first Table 1, which gives the bias of the various estimates as an estimate of $\alpha_{(N)}$. This is equivalent to the bias of estimated relative technical inefficiency $\hat{u}_i^*$ as an estimate of u_i^*. As expected, the bias of $\hat{\alpha}$ is larger when N is larger (the "max" is taken over more firms) and when μ_* is smaller (we are closer to a tie). The panel jackknife and the bias-corrected bootstrap are less biased than the fixed effects estimate $\hat{\alpha}$. However, they only correct part of the bias. In most cases the jackknife corrects more of the bias than the bias-corrected bootstrap. The generalized panel jackknife overcorrects (so the original upward bias now becomes a downward bias).

When μ_* is very small, so that the variability of the α_i is very small relative to the sampling variability of the $\hat{\alpha}_i$, we are in a sense close to a tie. In these cases the "exact tie" asymptotics appear to be relevant: the generalized panel jackknife is nearly unbiased, and the panel jackknife (and also the bias-corrected bootstrap) corrects about half of the bias, as predicted by Theorem 3. Conversely, when μ_* is large we are far from a tie, the panel jackknife and the bias-corrected bootstrap are nearly unbiased, and the downward bias of the generalized panel jackknife is almost as large as the upward bias of $\hat{\alpha}$, as predicted by Theorem 4.

Table 2 gives the variance of the various estimates. They are easy to summarize. The variance of the $\hat{\alpha}$ is less than the variance of the bias-corrected bootstrap point estimate, which is less than the variance of the panel jackknife, which is less than the variance of the generalized panel jackknife. The variance of the generalized panel jackknife is considerably larger than the variance of the other estimators. To properly interpret these variances, remember that we are ultimately interested in estimating the relative size of the u_i, whose variance is $(\sigma_u^2)_*$, and that in our setup $(\sigma_u^2)_* = 0.001, 0.01, 0.1, 1,$ and 10 for $\mu_* = 10^{-1}, 10^{-1/2}, 1, 10^{1/2},$ and 10, respectively. So the variance of these estimators is large enough to be an issue, except perhaps for the larger values of μ_*.

Table 3 gives the MSE of the estimates. In terms of MSE, the two varieties of the jackknife are dominated by the bias-corrected bootstrap. The bias-corrected bootstrap is also generally better than the fixed effects estimate $\hat{\alpha}$, except in those cases where the bias of $\hat{\alpha}$ is small (i.e., when N is small and μ_* is large).

Now we turn to Experiment II, the case of an exact tie. These results are in Table 4, 5, and 6.

In terms of bias, we see in Table 4 that the generalized panel jackknife is clearly the best. It overcorrects the bias, but not by as much as the panel jackknife and the bias-corrected bootstrap undercorrect. As expected from Theorem 3, the panel jackknife corrects about half of the bias. The bias-corrected bootstrap, which is not valid asymptotically in the case of an exact tie, also appears to correct about half of the bias.

In Table 5, the variances of the estimates are rather similar to the variances for the case of no tie (Table 2). The main difference is that now the variance does not depend as strongly on μ_*, presumably because, once we have forced a tie, the similarity of the other α_i is not of as much importance. The ranking of the estimators, in order of increasing variance, is still the same as in Table 2 ($\hat{\alpha}$, bias-corrected bootstrap, panel jackknife and generalized panel jackknife).

In terms of MSE, we see in Table 6 that the bias-corrected bootstrap still dominates both varieties of the jackknife. It is also generally better than the fixed effects estimate $\hat{\alpha}$. This favorable performance of the bias-corrected bootstrap is perhaps surprising, given that it is not asymptotically valid in the case of an exact tie.

Our last experiment is Experiment III, the case of a near tie. The results for this experiment are given in Table 7, 8, and 9. As a general statement, the results are between those of Experiment I and Experiment II, which is not surprising.

For small values of μ_* (nearer tie), the bias results in Table 7 are quite similar to those of Table 4 for an exact tie. In these cases the generalized panel jackknife has little bias, while the panel jackknife and the bias-corrected bootstrap correct about half of the bias. For large values of μ_* (less near tie), the panel jackknife and the bias-corrected bootstrap still correct only some of the bias, but the generalized panel jackknife overcorrects. Still, it is generally true in Table 7 that the generalized panel jackknife has the smallest bias.

In terms of variance (Table 8) and MSE (Table 9), the results are fairly similar to those for both the case of no tie and the case of an exact tie. Once again the bias-corrected bootstrap is generally the best, and the generalized panel jackknife is the worst.

The last issue we consider is the effect of changing T. We consider the same three kinds of experiments as just described, with $T = 5, 20, 50,$ and 100 (in addition to $T = 10$, which we have just discussed). These results are given in a set of 36 Supplemental Tables, available on request from the authors. In this paper we will display the results only for $\mu_* = 1$ and $N = 20$. Tables 10, 11, and 12 give the bias, variance, and MSE of the various estimates.

As discussed in Section 5, we do not expect changes in T to be very important, because we are holding constant N, μ_*, and σ_v^2/T, or equivalently we are holding constant N, $(\sigma_u^2)_*$, and σ_v^2/T. Indeed, the motivation for adopting this parameterization was that we expected it to make one of the parameters (T) unimportant. We expect changing T to be more important for the jackknife estimates than for the other two estimates, because the value of T affects the weights that the jackknife puts on the original estimate versus the average of the delete-one-observation estimates.

What we see in Tables 10 – 12 is not surprising. In Table 10, the effect of changing T on the bias of the estimates is very minor. In Table 11, changing T does not affect the variance of the fixed effects estimate or the bias-corrected bootstrap point estimate very much, but the variance of the jackknife estimates increases noticeably as T increases. Correspondingly, in Table 12 the MSE of the jackknife estimates increases as T increases. However, it remains true that the value of T is much less important than the values of N and μ_* in determining the relative performance of the various estimates.

7. CONCLUDING REMARKS

In the stochastic frontier model with panel data, the fixed effects estimate of the frontier intercept is biased upward. Previous work found that the bias-corrected bootstrap corrected only part of this bias. This paper has tried to explain that finding and to see whether we can more successfully remove the bias using the jackknife.

The bootstrap is known to be asymptotically (as $T \to \infty$ with N fixed) valid if there is no tie for the best firm, and not valid if there is an exact tie. So whether there is a tie, and how close we are to having a tie if there is not an exact tie, is a reasonable issue to focus on.

When there is an exact tie, we show that the bias of the fixed effects estimate is of order $T^{-1/2}$ rather than T^{-1}. Not only is the bootstrap not valid, but the usual panel jackknife, which is based on the assumption that the bias is of order T^{-1}, also does not work correctly. More specifically, we show that it removes (approximately) half of the bias. A different form of the jackknife, which we call the generalized panel jackknife, is needed to remove the bias of order $T^{-1/2}$.

If there is no tie, the bootstrap is valid and the panel jackknife should also be effective in removing bias, since now the bias is of order T^{-1}. In this case the generalized panel jackknife will not work correctly, and indeed we show that its bias is the negative of the bias of the fixed effects estimate; it reverses the bias.

We also consider the case of a near tie, which we define as the case that the difference between the frontier intercept and the intercept of the second-best firm is $O(T^{-1/2})$. In this case the bias is again of order $T^{-1/2}$ and so the generalized panel jackknife should remove it.

Our simulations support the finite-sample relevance of these arguments. When there is a tie or a near tie, the generalized panel jackknife removes the bias effectively, whereas the panel jackknife and the bias-corrected bootstrap remove about half of the bias. When there is not a tie, the generalized panel jackknife overcorrects the bias, and the panel jackknife and the bias-corrected bootstrap are much better at removing the bias.

The major drawback of the jackknife is that its variance is large. This is true for both versions of the jackknife but the variance is the largest for the generalized panel jackknife. There does not seem to be any good reason to prefer the panel jackknife to the bias-corrected bootstrap, since it has a larger variance and does not do a better job of correcting bias. However, while the generalized panel jackknife is clearly dominated by the bias-corrected bootstrap in terms of MSE, it does do a very good job of removing bias when there is an exact tie or a near tie. Empirically, presumably that corresponds to cases where the identity of the best firm is in substantial doubt.

The inability of the generalized panel jackknife to beat the bias-corrected bootstrap in terms of MSE when there is an exact or a near tie is perhaps surprising, since the bootstrap is not valid if there is a tie. However, “not valid” here has a specific meaning, namely that we cannot claim that the distribution of the bootstrap estimate around the original estimate matches the distribution of the original estimate around the true parameter. Apparently the bias-corrected bootstrap is nevertheless a useful point estimate.

REFERENCES

Dhaene, G., K. Jochmans, and B. Thuysbaert, 2006, Split-Panel Jackknife Estimation of Fixed Effects Models (Previous Title: Jackknife Bias Reduction for Nonlinear Dynamic Panel Data Models with Fixed Effects), Working Paper.

Fernández-Val, I. and F. Vella, 2007, Bias Corrections for Two-Step Fixed Effects Panel Data Estimator, IZA Discussion Paper 2690, Institute for the Study of Labor (IZA)

Gray, H.L., and W.R. Schucany, 1972, *The Generalized Jackknife Statistic*, Marcel Dekker, Inc., New York.

Hahn, J. and G. Kuersteiner, 2004, Bias Reduction for Dynamic Nonlinear Panel Models with Fixed Effects, Unpublished Manuscript.

Hahn, J., and W. Newey, 2004, Jackknife and Analytical Bias Reduction for Nonlinear Panel Models, *Econometrica*, 72, 1295-1319.

Hall, P., W.Härdle, and L. Simar, 1995, Iterated Bootstrap with Applications to Frontier Models, *Journal of Productivity Analysis*, 6, 63-76.

Kim, M., Y. Kim, and P. Schmidt, 2007, On the Accuracy of Bootstrap Confidence Intervals for Efficiency Levels in Stochastic Frontier Models with Panel Data, *Journal of Productivity Analysis*, 28, 165-181.

Miller, R.G., 1974, The Jackknife – A Review, *Biometrika*, 61, 1-15.

Park, B.U., and L. Simar, 1994, Efficient Semiparametric Estimation in a Stochastic Frontier Model, *Journal of the American Statistics Association*, 89, 929-936.

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

Quenouille, M.H., 1956, Note on Bias in Estimation, *Biometrika*, 61, 353-360.

Schmidt, P., and R. Sickles, 1984, Production Frontiers and Panel Data, *Journal of Business and Economic Statistics*, 2, 367-374.

Schucany, W.R., H.L. Gray, and D.B. Owen, 1971, On Bias Reduction in Estimation, *Journal of the American Statistical Association*, 66, 524-533.

Tukey, J.W., 1958, Bias and Confidence in Not Quite Large Samples, (Abstract), *Annals of Mathematical Statistics*, 18, 614.

TABLE 1

EXPERIMENT I: NO TIE

$T = 10$

BIAS OF THE ESTIMATES

μ_*	N	(1) $E\left(\hat{\alpha} - \alpha_{(N)}\right)$	(2) $E\left(J(\hat{\alpha}) - \alpha_{(N)}\right)$	(3) $E\left(G(\hat{\alpha}) - \alpha_{(N)}\right)$	(4) $E\left(\hat{\alpha}_{BC}^{boot} - \alpha_{(N)}\right)$
10^{-1}	2	0.1671	0.0810	-0.0099	0.1006
$10^{-1/2}$	2	0.1391	0.0522	-0.0394	0.0743
1	2	0.0887	0.0230	-0.0463	0.0368
$10^{1/2}$	2	0.0462	0.0125	-0.0230	0.0204
10	2	0.0294	0.0199	0.0100	0.0204
10^{-1}	10	0.4532	0.2113	-0.0436	0.2669
$10^{-1/2}$	10	0.3935	0.1556	-0.0951	0.2111
1	10	0.2809	0.0828	-0.1261	0.1218
$10^{1/2}$	10	0.1504	0.0293	-0.0983	0.0439
10	10	0.0577	-0.0034	-0.0678	0.0046
10^{-1}	20	0.5566	0.2724	-0.0271	0.3326
$10^{-1/2}$	20	0.4928	0.2093	-0.0895	0.2722
1	20	0.3750	0.1176	-0.1537	0.1767
$10^{1/2}$	20	0.2349	0.0563	-0.1321	0.0895
10	20	0.1136	0.0074	-0.1046	0.0292
10^{-1}	50	0.6699	0.3132	-0.0629	0.3975
$10^{-1/2}$	50	0.6092	0.2678	-0.0921	0.3433
1	50	0.4973	0.1903	-0.1334	0.2565
$10^{1/2}$	50	0.3556	0.1151	-0.1385	0.1639
10	50	0.2059	0.0379	-0.1391	0.0724
10^{-1}	100	0.7584	0.3627	-0.0544	0.4594
$10^{-1/2}$	100	0.6949	0.3127	-0.0901	0.4012
1	100	0.5809	0.2293	-0.1413	0.3086
$10^{1/2}$	100	0.4433	0.1652	-0.1280	0.2182
10	100	0.2950	0.0922	-0.1217	0.1281

TABLE 2

EXPERIMENT I: NO TIE

$T = 10$

VARIANCE OF THE ESTIMATES

μ_*	N	(1) $\hat{\alpha}$	(2) $J(\hat{\alpha})$	(3) $G(\hat{\alpha})$	(4) $\hat{\alpha}_{BC}^{boot}$
10^{-1}	2	0.0666	0.1130	0.2127	0.0805
$10^{-1/2}$	2	0.0696	0.1159	0.2149	0.0839
1	2	0.0803	0.1183	0.2002	0.0936
$10^{1/2}$	2	0.0932	0.1133	0.1582	0.1004
10	2	0.0991	0.1055	0.1189	0.1018
10^{-1}	10	0.0355	0.1440	0.3871	0.0623
$10^{-1/2}$	10	0.0382	0.1472	0.3901	0.0662
1	10	0.0483	0.1478	0.3665	0.0764
$10^{1/2}$	10	0.0696	0.1378	0.2836	0.0938
10	10	0.0890	0.1282	0.2106	0.1034
10^{-1}	20	0.0264	0.1419	0.4089	0.0528
$10^{-1/2}$	20	0.0284	0.1467	0.4191	0.0557
1	20	0.0359	0.1534	0.4191	0.0650
$10^{1/2}$	20	0.0518	0.1388	0.3197	0.0775
10	20	0.0757	0.1329	0.2613	0.0948
10^{-1}	50	0.0208	0.1500	0.4570	0.0469
$10^{-1/2}$	50	0.0215	0.1489	0.4518	0.0476
1	50	0.0254	0.1479	0.4356	0.0527
$10^{1/2}$	50	0.0359	0.1438	0.3896	0.0638
10	50	0.0569	0.1401	0.3257	0.0832
10^{-1}	100	0.0191	0.1617	0.5014	0.0466
$10^{-1/2}$	100	0.0196	0.1556	0.4799	0.0467
1	100	0.0231	0.1587	0.4807	0.0515
$10^{1/2}$	100	0.0317	0.1585	0.4490	0.0626
10	100	0.0440	0.1400	0.3532	0.0716

TABLE 3

EXPERIMENT I: NO TIE

$T = 10$

MSE OF THE ESTIMATES

μ_*	N	(1) $\hat{\alpha}$	(2) $J(\hat{\alpha})$	(3) $G(\hat{\alpha})$	(4) $\hat{\alpha}_{BC}^{boot}$
10^{-1}	2	0.0945	0.1196	0.2127	0.0906
$10^{-1/2}$	2	0.0890	0.1186	0.2164	0.0894
1	2	0.0882	0.1188	0.2024	0.0950
$10^{1/2}$	2	0.0953	0.1134	0.1588	0.1008
10	2	0.1000	0.1059	0.1190	0.1022
10^{-1}	10	0.2409	0.1887	0.3891	0.1355
$10^{-1/2}$	10	0.1931	0.1715	0.3991	0.1107
1	10	0.1272	0.1547	0.3824	0.0912
$10^{1/2}$	10	0.0922	0.1386	0.2932	0.0958
10	10	0.0923	0.1282	0.2152	0.1035
10^{-1}	20	0.3362	0.2162	0.4096	0.1634
$10^{-1/2}$	20	0.2712	0.1905	0.4271	0.1298
1	20	0.1765	0.1673	0.4427	0.0962
$10^{1/2}$	20	0.1070	0.1420	0.3472	0.0855
10	20	0.0886	0.1329	0.2722	0.0956
10^{-1}	50	0.4696	0.2481	0.4610	0.2049
$10^{-1/2}$	50	0.3925	0.2206	0.4603	0.1655
1	50	0.2727	0.1841	0.4534	0.1184
$10^{1/2}$	50	0.1623	0.1571	0.4087	0.0907
10	50	0.0993	0.1415	0.3451	0.0885
10^{-1}	100	0.5942	0.2932	0.5044	0.2576
$10^{-1/2}$	100	0.5025	0.2534	0.4880	0.2077
1	100	0.3605	0.2112	0.5006	0.1467
$10^{1/2}$	100	0.2282	0.1858	0.4654	0.1102
10	100	0.1310	0.1485	0.3680	0.0880

TABLE 4

EXPERIMENT II: EXACT TIE

$T = 10$

BIAS OF THE ESTIMATES

μ_*	N	(1) $E(\hat{\alpha} - \alpha_{(N)})$	(2) $E(J(\hat{\alpha}) - \alpha_{(N)})$	(3) $E(G(\hat{\alpha}) - \alpha_{(N)})$	(4) $E(\hat{\alpha}_{BC}^{boot} - \alpha_{(N)})$
**	2	0.1828	0.0997	0.0120	0.1163
10^{-1}	10	0.4580	0.2165	-0.0380	0.2724
$10^{-1/2}$	10	0.4077	0.1693	-0.0820	0.2256
1	10	0.3261	0.1268	-0.0832	0.1678
$10^{1/2}$	10	0.2499	0.1122	-0.0330	0.1323
10	10	0.2052	0.0817	-0.0424	0.1145
10^{-1}	20	0.5593	0.2696	-0.0357	0.3353
$10^{-1/2}$	20	0.5020	0.2158	-0.0859	0.2827
1	20	0.3988	0.1412	-0.1302	0.2006
$10^{1/2}$	20	0.2938	0.0976	-0.1093	0.1421
10	20	0.2282	0.0954	-0.0446	0.1194
10^{-1}	50	0.6707	0.3104	-0.0693	0.3985
$10^{-1/2}$	50	0.6134	0.2715	-0.0889	0.3492
1	50	0.5124	0.2117	-0.1052	0.2739
$10^{1/2}$	50	0.3926	0.1579	-0.0895	0.2038
10	50	0.2899	0.1237	-0.0514	0.1530
10^{-1}	100	0.7615	0.3702	-0.0422	0.4642
$10^{-1/2}$	100	0.7023	0.3240	-0.0747	0.4117
1	100	0.5950	0.2523	-0.1089	0.3271
$10^{1/2}$	100	0.4665	0.1843	-0.1132	0.2434
10	100	0.3382	0.1236	-0.1026	0.1660

Note: ** value of μ_* is irrelevant when $N = 2$ and there is an exact tie.

TABLE 5

EXPERIMENT II: EXACT TIE

$T = 10$

VARIANCE OF THE ESTIMATES

μ_*	N	(1) $\hat{\alpha}$	(2) $J(\hat{\alpha})$	(3) $G(\hat{\alpha})$	(4) $\hat{\alpha}_{BC}^{boot}$
**	2	0.0658	0.1107	0.2063	0.0792
10^{-1}	10	0.0346	0.1408	0.3785	0.0607
$10^{-1/2}$	10	0.0362	0.1411	0.3785	0.0626
1	10	0.0424	0.1396	0.3598	0.0682
$10^{1/2}$	10	0.0528	0.1210	0.2729	0.0748
10	10	0.0620	0.1308	0.2785	0.0813
10^{-1}	20	0.0264	0.1449	0.4207	0.0531
$10^{-1/2}$	20	0.0282	0.1464	0.4187	0.0556
1	20	0.0351	0.1484	0.4048	0.0640
$10^{1/2}$	20	0.0471	0.1444	0.3596	0.0746
10	20	0.0562	0.1228	0.2694	0.0778
10^{-1}	50	0.0208	0.1495	0.4570	0.0474
$10^{-1/2}$	50	0.0225	0.1554	0.4721	0.0506
1	50	0.0274	0.1554	0.4506	0.0579
$10^{1/2}$	50	0.0369	0.1505	0.4032	0.0685
10	50	0.0497	0.1385	0.3311	0.0765
10^{-1}	100	0.0192	0.1614	0.5011	0.0471
$10^{-1/2}$	100	0.0204	0.1594	0.4892	0.0492
1	100	0.0247	0.1606	0.4784	0.0556
$10^{1/2}$	100	0.0313	0.1551	0.4395	0.0620
10	100	0.0430	0.1443	0.3680	0.0715

Note: ** value of μ_* is irrelevant when $N = 2$ and there is an exact tie.

TABLE 6

EXPERIMENT II: EXACT TIE

$T = 10$

MSE OF THE ESTIMATES

μ_*	N	(1) $\hat{\alpha}$	(2) $J(\hat{\alpha})$	(3) $G(\hat{\alpha})$	(4) $\hat{\alpha}_{BC}^{boot}$
**	2	0.0993	0.1207	0.2065	0.0928
10^{-1}	10	0.2444	0.1877	0.3811	0.1350
$10^{-1/2}$	10	0.2024	0.1698	0.3852	0.1135
1	10	0.1437	0.1556	0.3667	0.0964
$10^{1/2}$	10	0.1153	0.1336	0.2740	0.0923
10	10	0.1041	0.1379	0.2803	0.0944
10^{-1}	20	0.3392	0.2176	0.4220	0.1655
$10^{-1/2}$	20	0.2802	0.1929	0.4260	0.1355
1	20	0.1941	0.1683	0.4218	0.1043
$10^{1/2}$	20	0.1334	0.1539	0.3716	0.0948
10	20	0.1083	0.1319	0.2714	0.0920
10^{-1}	50	0.4706	0.2459	0.4618	0.2063
$10^{-1/2}$	50	0.3987	0.2291	0.4800	0.1726
1	50	0.2900	0.2002	0.4617	0.1340
$10^{1/2}$	50	0.1910	0.1754	0.4112	0.1101
10	50	0.1337	0.1538	0.3337	0.0999
10^{-1}	100	0.5991	0.2984	0.5029	0.2626
$10^{-1/2}$	100	0.5136	0.2644	0.4948	0.2187
1	100	0.3787	0.2243	0.4903	0.1626
$10^{1/2}$	100	0.2489	0.1801	0.4523	0.1212
10	100	0.1573	0.1596	0.3785	0.0991

Note: ** value of μ_* is irrelevant when $N = 2$ and there is an exact tie.

TABLE 7

EXPERIMENT III: NEAR TIE

$T = 10$

BIAS OF THE ESTIMATES

μ_*	N	(1) $E(\hat{\alpha} - \alpha_{(N)})$	(2) $E(J(\hat{\alpha}) - \alpha_{(N)})$	(3) $E(G(\hat{\alpha}) - \alpha_{(N)})$	(4) $E(\hat{\alpha}_{BC}^{boot} - \alpha_{(N)})$
10^{-1}	2	0.1774	0.0934	0.0049	0.1107
$10^{-1/2}$	2	0.1661	0.0807	-0.0094	0.0994
1	2	0.1361	0.0503	-0.0402	0.0707
$10^{1/2}$	2	0.0792	0.0107	-0.0616	0.0252
10	2	0.0323	0.0003	-0.0334	0.0053
10^{-1}	10	0.4578	0.2164	-0.0381	0.2724
$10^{-1/2}$	10	0.4067	0.1678	-0.0841	0.2244
1	10	0.3210	0.1207	-0.0904	0.1620
$10^{1/2}$	10	0.2273	0.0861	-0.0627	0.1084
10	10	0.1403	0.0326	-0.0809	0.0560
10^{-1}	20	0.5592	0.2695	-0.0358	0.3353
$10^{-1/2}$	20	0.5018	0.2158	-0.0857	0.2826
1	20	0.3971	0.1390	-0.1331	0.1986
$10^{1/2}$	20	0.2844	0.0851	-0.1249	0.1314
10	20	0.1906	0.0534	-0.0912	0.0808
10^{-1}	50	0.6707	0.3104	-0.0693	0.3985
$10^{-1/2}$	50	0.6133	0.2713	-0.0891	0.3491
1	50	0.5119	0.2112	-0.1059	0.2754
$10^{1/2}$	50	0.3898	0.1537	-0.0952	0.2004
10	50	0.2755	0.1053	-0.0741	0.1367
10^{-1}	100	0.7615	0.3702	-0.0422	0.4642
$10^{-1/2}$	100	0.7023	0.3240	-0.0748	0.4117
1	100	0.5950	0.2522	-0.1090	0.3270
$10^{1/2}$	100	0.4659	0.1835	-0.1141	0.2426
10	100	0.3340	0.1190	-0.1075	0.1613

TABLE 8

EXPERIMENT III: NEAR TIE

$T = 10$

VARIANCE OF THE ESTIMATES

μ_*	N	(1) $\hat{\alpha}$	(2) $J(\hat{\alpha})$	(3) $G(\hat{\alpha})$	(4) $\hat{\alpha}_{BC}^{boot}$
10^{-1}	2	0.0660	0.1117	0.2092	0.0795
$10^{-1/2}$	2	0.0666	0.1136	0.2137	0.0804
1	2	0.0696	0.1193	0.2251	0.0840
$10^{1/2}$	2	0.0820	0.1234	0.2086	0.0963
10	2	0.0956	0.1174	0.1622	0.1030
10^{-1}	10	0.0347	0.1408	0.3794	0.0607
$10^{-1/2}$	10	0.0363	0.1412	0.3786	0.0626
1	10	0.0426	0.1396	0.3605	0.0683
$10^{1/2}$	10	0.0548	0.1250	0.2817	0.0762
10	10	0.0686	0.1243	0.2539	0.0840
10^{-1}	20	0.0264	0.1449	0.4206	0.0531
$10^{-1/2}$	20	0.0282	0.1464	0.4188	0.0556
1	20	0.0351	0.1483	0.4050	0.0639
$10^{1/2}$	20	0.0475	0.1468	0.3663	0.0750
10	20	0.0597	0.1277	0.2794	0.0804
10^{-1}	50	0.0208	0.1495	0.4570	0.0474
$10^{-1/2}$	50	0.0225	0.1554	0.4772	0.0506
1	50	0.0274	0.1552	0.4502	0.0578
$10^{1/2}$	50	0.0368	0.1517	0.4075	0.0684
10	50	0.0504	0.1403	0.3369	0.0771
10^{-1}	100	0.0192	0.1612	0.5011	0.0471
$10^{-1/2}$	100	0.0204	0.1594	0.4893	0.0492
1	100	0.0247	0.1606	0.4784	0.0556
$10^{1/2}$	100	0.0313	0.1548	0.4384	0.0619
10	100	0.0430	0.1440	0.3674	0.0715

TABLE 9

EXPERIMENT III: NEAR TIE

$T = 10$

MSE OF THE ESTIMATES

μ_*	N	(1) $\hat{\alpha}$	(2) $J(\hat{\alpha})$	(3) $G(\hat{\alpha})$	(4) $\hat{\alpha}_{BC}^{boot}$
10^{-1}	2	0.0975	0.1204	0.2093	0.0918
$10^{-1/2}$	2	0.0942	0.1201	0.2138	0.0903
1	2	0.0881	0.1218	0.2267	0.0890
$10^{1/2}$	2	0.0883	0.1235	0.2124	0.0969
10	2	0.0966	0.1174	0.1633	0.1031
10^{-1}	10	0.2442	0.1876	0.3809	0.1349
$10^{-1/2}$	10	0.2016	0.1694	0.3857	0.1130
1	10	0.1456	0.1542	0.3686	0.0945
$10^{1/2}$	10	0.1056	0.1324	0.2856	0.0880
10	10	0.0883	0.1254	0.2605	0.0871
10^{-1}	20	0.3391	0.2175	0.4219	0.1655
$10^{-1/2}$	20	0.2800	0.1930	0.4261	0.1354
1	20	0.1928	0.1676	0.4227	0.1034
$10^{1/2}$	20	0.1284	0.1540	0.3819	0.0922
10	20	0.0960	0.1305	0.2877	0.0870
10^{-1}	50	0.4706	0.2459	0.4618	0.2062
$10^{-1/2}$	50	0.3986	0.2290	0.4801	0.1725
1	50	0.2895	0.1998	0.4614	0.1336
$10^{1/2}$	50	0.1887	0.1753	0.4166	0.1085
10	50	0.1263	0.1514	0.3424	0.0958
10^{-1}	100	0.5991	0.2984	0.5029	0.2626
$10^{-1/2}$	100	0.5136	0.2644	0.4949	0.2187
1	100	0.3786	0.2242	0.4903	0.1626
$10^{1/2}$	100	0.2483	0.1885	0.4514	0.1208
10	100	0.1545	0.1582	0.3789	0.0975

TABLE 10

EFFECT OF CHANGING T

$\mu_* = 1, N = 20$

BIAS OF THE ESTIMATES

Experiment	T	(1) $E(\hat{\alpha} - \alpha_{(N)})$	(2) $E(J(\hat{\alpha}) - \alpha_{(N)})$	(3) $E(G(\hat{\alpha}) - \alpha_{(N)})$	(4) $E(\hat{\alpha}_{BC}^{boot} - \alpha_{(N)})$
I NO TIE	5	0.3842	0.1333	-0.1472	0.2098
	10	0.3750	0.1176	-0.1537	0.1767
	20	0.3795	0.1417	-0.1023	0.1782
	50	0.3821	0.1380	-0.1085	0.1750
	100	0.3764	0.1292	-0.1193	0.1659
II EXACT TIE	5	0.4032	0.1566	-0.1191	0.2272
	10	0.3988	0.1412	-0.1302	0.2066
	20	0.3908	0.1407	-0.1220	0.1900
	50	0.4037	0.1556	-0.0950	0.1969
	100	0.3979	0.1635	-0.0720	0.1877
III NEAR TIE	5	0.4012	0.1547	-0.1210	0.2252
	10	0.3971	0.1390	-0.1331	0.1986
	20	0.3950	0.1389	-0.1242	0.1885
	50	0.4031	0.1563	-0.0930	0.1962
	100	0.3973	0.1614	-0.0757	0.1871

TABLE 11

EFFECT OF CHANGING T

$\mu_* = 1, N = 20$

VARIANCE OF THE ESTIMATES

Experiment	T	(1) $\hat{\alpha}$	(2) $J(\hat{\alpha})$	(3) $G(\hat{\alpha})$	(4) $\hat{\alpha}_{BC}^{boot}$
I NO TIE	5	0.0365	0.1208	0.3162	0.0647
	10	0.0359	0.1534	0.4191	0.0650
	20	0.0365	0.1678	0.4717	0.0669
	50	0.0363	0.2247	0.7607	0.0663
	100	0.0378	0.3265	1.0631	0.0702
II EXACT TIE	5	0.0334	0.1139	0.3029	0.0593
	10	0.0351	0.1484	0.4048	0.0640
	20	0.0361	0.1867	0.5337	0.0650
	50	0.0384	0.2568	0.7893	0.0714
	100	0.0366	0.3046	0.9902	0.0680
III NEAR TIE	5	0.0332	0.1137	0.3030	0.0590
	10	0.0351	0.1483	0.4050	0.0639
	20	0.0360	0.1868	0.5343	0.0649
	50	0.0384	0.2542	0.7810	0.0713
	100	0.0366	0.3089	1.0055	0.0680

TABLE 12

EFFECT OF CHANGING T

$\mu_* = 1, N = 20$

MSE OF THE ESTIMATES

Experiment	T	(1) $\hat{\alpha}$	(2) $J(\hat{\alpha})$	(3) $G(\hat{\alpha})$	(4) $\hat{\alpha}_{BC}^{boot}$
I NO TIE	5	0.1841	0.1386	0.3379	0.1087
	10	0.1765	0.1673	0.4427	0.0962
	20	0.1805	0.1879	0.4822	0.0987
	50	0.1824	0.2637	0.7725	0.0969
	100	0.1795	0.3432	1.0773	0.0977
II EXACT TIE	5	0.1959	0.1384	0.3171	0.1109
	10	0.1941	0.1683	0.4218	0.1043
	20	0.1935	0.2065	0.5486	0.1011
	50	0.2014	0.2810	0.7989	0.1102
	100	0.1949	0.3313	0.9954	0.1033
III NEAR TIE	5	0.1942	0.1376	0.3176	0.1097
	10	0.1928	0.1676	0.4227	0.1034
	20	0.1924	0.2061	0.5497	0.1004
	50	0.2009	0.2787	0.7896	0.1098
	100	0.1945	0.3349	1.0113	0.1030