L
2009

This is to certify that the
dissertation entitled

THREE ESSAYS IN ECONOMETRICS

presented by

PANUTAT SATCHACHAI

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in Economics

Peter Schmidt

Major Professor's Signature

JULY 22, 2009

Date

THREE ESSAYS IN ECONOMETRICS

By

Panutat Satchachai

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILSOPHY

Economics

2009

ABSTRACT

THREE ESSAYS IN ECONOMETRICS

By

Panutat Satchachai

In the first chapter, we consider GMM estimation when there are more moment conditions than observations. Due to the singularity of the estimated variance matrix of the moment conditions, the quick solution of using the generalized inverse, although temping, is shown to be unfruitful.

In the second and third chapters, we consider the problem of point estimation of technical inefficiency in a simple stochastic frontier model with panel data.

In the second chapter, we wish to correct the bias of the estimates of technical inefficiency based on fixed effects estimation that previously shown to be biased upward. Previous work has attempted to correct this bias using the bootstrap, but in simulations the bootstrap correct 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 successfully in removing the bias. However, the resulting estimates have a large variance.

In the third chapter, we focus on how we could decrease the variance and MSE of a jackknife-type estimate of the frontier intercept found in the previous chapter. We consider the split-sample jackknife proposed by Dhaene, Jochmans and Thuysbaert (2006), which is simply two times the original estimate based on the whole sample minus

the average of the two half-sample estimates, and the "generalized" version proposed by Satchachai and Schmidt (2008), which is relevant in the case of an exact tie or a near tie. Although these estimators also successfully remove the bias, their variance is still large. We also consider whether or not there is an "optimal" split-sample jackknife estimator that has small variance and/or small MSE. For a special case of $N = 2$, we derive the "optimal" weights for the original estimate and the half-sample estimates. Although the "optimal" split-sample jackknife has even smaller variance and MSE, it does not properly remove the bias, and it appears that there is not much gain in terms of mean square error from applying the jackknife procedure.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Professor Peter Schmidt. Without his guidance and encouragement, this thesis would have not appeared in its finished form. His teaching approach, research experience and knowledge in economics also help me developing my teaching method as well as training and shaping the direction of my research interests and of this dissertation. His help in editing the dissertation is also gratefully acknowledged. I would also like to thank Assistant Professor Emma Iglesias who was always willing to spend time in clarifying some doubts encountered and gave invaluable feedback during the process of writing this dissertation, and to Professor Timothy Vogelsang for his invaluable comments. I am also grateful for the teaching opportunity given by the Department of Economics through the graduate assistantship.

I am also thankful of my undergraduate friends, fellow graduate students and the staff at the Department of Economics who also contributed to the completion of this dissertation through their friendship and help. A special thank you also goes to my friends at the Michigan State University badminton club and the Thai student community at Michigan State University for their warm-hearted friendship.

Finally, I am greatly indebted to and thankful of my parents for their support, both financially and mentally. My special thanks also go to my sisters for their love and encouragement and to my grandmother who was always supportive of my decision. I could only wish that she could share this joyful moment with me.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER ONE

Satchachai, Panutat and Peter Schmidt, 2008, GMM with More Moment Conditions than Observations, *Economics Letters*, 99, 252-255.

Chapter 1

GMM with More Moment Conditions than Observations

1.1 Introduction

In this paper, we consider GMM estimation when there are more moment conditions than observations. This can occur in practice, for example, in dynamic panel data models, as in Han et al. (2005). It is well known that the optimal weighting matrix for GMM is the inverse of the variance matrix of the moment conditions. However, when there are more moment conditions than observations, the usual estimate of the variance matrix of the moment conditions is singular. In that case it is tempting to use its generalized inverse as the weighting matrix. The point of this paper is to demonstrate that this is not a good idea. When the continuous updating form of GMM is used, the value of the criterion function equals one for all values of the parameter. When the two step GMM estimate is used, the value of the criterion function is less than or equal to one, and again the usefulness of such procedure is doubtful.

1.2 The Model

We consider GMM estimation based on the population moment conditions

$$E(g(y,\theta_0)) = 0 . \tag{1.1}$$

Here g is a $k \times 1$ vector of moment conditions, and θ_0 is the population value of a p - dimensional parameter θ. We assume $k \geq p$ so there are enough moment conditions to identify θ_0.

The data $y_1,...,y_n$ are a random sample from a population that satisfies (1.1). In the usual case, $n > k$, but in this paper we consider cases where $n < k$.

We define the sample moment conditions as

$$\bar{g}_n = \frac{1}{n}\sum_{i=1}^{n} g(y_i,\theta). \tag{1.2}$$

Let $C = Eg(y_i,\theta)g(y_i,\theta)'$, the variance matrix of the population moment conditions. The usual estimate of C is

$$\hat{C}_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} g(y_i,\theta)g(y_i,\theta)'. \tag{1.3}$$

Note that $\hat{C}_n(\theta)$ is singular when $n < k$.

We now distinguish two different forms of the GMM estimator. The "continuous updating" estimator $\hat{\theta}_{CUE}$ is the value of θ that minimizes the criterion function

$$Q_n^{CUE}(\theta) = \bar{g}_n(\theta)'\hat{C}_n(\theta)^{-1}\bar{g}_n(\theta). \tag{1.4}$$

The "two step" estimator $\hat{\theta}_{2STEP}$ is the value of θ that minimizes the criterion function

$$Q_n^{2STEP}(\theta) = \bar{g}_n(\theta)'\hat{C}_n(\hat{\theta})^{-1}\bar{g}_n(\theta). \tag{1.5}$$

where $\hat{\theta}$ is some initial (consistent) estimator.

When $n < k$, these estimates do not exist. However, we can replace the inverse $\hat{C}_n(\theta)^{-1}$ and $\hat{C}_n(\hat{\theta})^{-1}$ by the "generalized inverse" $\hat{C}_n(\theta)^{+}$ and $\hat{C}_n(\hat{\theta})^{+}$, in (1.4) and (1.5) respectively.[1] For example, Han, Orea and Schmidt (2005) follow this procedure.

[1] If for any $m \times n$ matrix H with rank of r, then it is always possible to find matrices R, a $m \times r$ matrix, and S, each having a rank of r, such that $H = RS$. Then the generalized inverse of H, H^{+}, is

1. 3 The Case of $n=1$

Suppose that $n=1$, so $\bar{g}_n(\theta)=g(y_1,\theta)$.

We note that, if A is a $k\times k$ matrix of rank 1, so that $A=\xi\xi'$, where ξ is a $k\times 1$, then $A^+=(\xi'\xi)^{-2}A$.

Result 1.1 (i) $Q_n^{CUE}(\theta)=1$ for all θ.

(ii) $Q_n^{2STEP}(\hat{\theta}_{2STEP})\le 1$.

Proof. (i) $Q_n^{CUE}(\theta)=g(y_1,\theta)'[g(y_1,\theta)g(y_1,\theta)']^{+}g(y_1,\theta)$

$$=g(y_1,\theta)'g(y_1,\theta)[g(y_1,\theta)g(y_1,\theta)']^{-2}g(y_1,\theta)'g(y_1,\theta)$$
$$=1.$$

(ii) $Q_n^{2STEP}(\hat{\theta}_{2STEP})\le Q_n^{2STEP}(\hat{\theta})=1$ for an arbitrary $\hat{\theta}$. □

Part (i) of this result says that $\hat{\theta}_{CUE}$ is not defined, because the criterion function $Q_n^{CUE}(\theta)$ equals one for all θ. Clearly this criterion function contains no information at all about θ.

Part (ii) of this result says that the minimized value of $Q_n^{2STEP}(\hat{\theta}_{2STEP})$ is bounded above by one. It may or may not equal one.

$$H^+=(RS)^+=S'(SS')^{-1}(R'R)^{-1}R'.$$

The generalized inverse (or "Moore-Penrose pseudo inverse") H^+ is the unique matrix that satisfies the following properties

(i) HH^+ and H^+H are symmetric;

(ii) $H^+HH^+=H^+$; and

(iii) $HH^+H=H$.

If H is invertible, then $H^+=H^{-1}$.

Example 1.1 Suppose that y is $k \times 1$, and every element of y has expectation equal to μ. Thus we have

$$E(y - \mu e_k) = 0 \tag{1.6}$$

where e_k is a $k \times 1$ vector of ones. We are not restricting the variance matrix of y. We have one observation, y_1 $(k \times 1)$. Note that an obvious estimator is $\hat{\mu} = \frac{1}{k} e_k' y_1$, and this would have certain optimality properties if the elements of y are uncorrelated and have equal variance, but not otherwise.

By part (i) of Result 1.1, the continuous updating estimate of μ is undefined. We have $(y - \mu e_k)' [(y - \mu e_k)(y - \mu e_k)']^+ (y - \mu e_k) = 1$ for all μ.

An interesting result that holds for this example is that the two step estimate based on the initial estimate $\hat{\mu}$ (as above) is also undefined. We have

$$(y - \mu e_k)' [(y - \hat{\mu} e_k)(y - \hat{\mu} e_k)']^+ (y - \mu e_k) = \frac{[(y - \mu e_k)'(y - \hat{\mu} e_k)]^2}{[(y - \hat{\mu} e_k)'(y - \hat{\mu} e_k)]^2}.$$

and this equals one for all μ, because $e_k'(y - \hat{\mu} e_k) = 0$ for this specific choice of $\hat{\mu}$.

1.4 The General Case

The problem is exactly as Section 1.2. We define $\bar{g}_n(\theta)$ as in equation (1.2), and the criterion function $Q_n^{CUE}(\theta)$ and $Q_n^{2STEP}(\theta)$ as in equation (1.4) and (1.5). We have n observations, with $n < k$.

One basic result is essentially the same as for the case of $n = 1$.

Result 1.2 (i) $Q_n^{CUE}(\theta)=1$ for all θ.

(ii) $Q_n^{2STEP}(\hat{\theta}_{2STEP})\leq 1$.

Proof. See Appendix. □

So, as in the case of $n=1$, the CUE criterion function contains no information about θ. Similarly $Q_n^{2STEP}(\theta)$ is bounded above by one, and is not of any obvious use.

Unlike the case of $n=1$, it does not seem generally possible to say anything useful about circumstance in which $Q_n^{2STEP}(\hat{\theta}_{2STEP})=1$. An exception is the following example.

Example 1.2 This is the same as Example 1.1 except that now $n=2$. So we observe y_1 and y_2, where $E(y_i)=\mu e_k$ and $V(y_i)=\Sigma$, an unrestricted $k\times k$ matrix, for $i=1,2$. Define $\bar{y}=\frac{1}{2}(y_1+y_2)$, which is $k\times 1$, and define the scalar $\bar{\bar{y}}=\frac{1}{k}e_k'\bar{y}$ $=\frac{1}{2k}\sum_{i=1}^{2}\sum_{j=1}^{k}y_{ij}$. Let $\tilde{\mu}$ be the two step estimator based on the initial estimator $\bar{\bar{y}}$. That is, $\tilde{\mu}$ minimizes the criterion function

$$Q_n^{2STEP}(\mu)=(\bar{y}-\mu e_k)'\left[\frac{1}{2}\sum_{i=1}^{2}(y_i-\bar{\bar{y}}e_k)(y_i-\bar{\bar{y}}e_k)'\right]^{+}(\bar{y}-\mu e_k). \quad (1.7)$$

From Result 1.2 we know that $Q_n^{2STEP}(\tilde{\mu})\leq 1$. The following result (provided in the Appendix) gives the condition for equality.

Result 1.3 In Example 1.2, we have $\tilde{\mu} = \bar{\bar{y}}$ and $Q_n^{2STEP}(\tilde{\mu}) = 1$, if

$$\sum_{j=1}^{k}(y_{1j} - \bar{y}_1)^2 = \sum_{j=1}^{k}(y_{2j} - \bar{y}_2)^2. \tag{1.8}$$

1.5 Concluding Remarks

When there are more moment conditions than observations, the usual estimate of the variance matrix of the moment conditions is singular, and so the usual "optimal" weighting matrix cannot be calculated. This paper shows that this problem cannot be solved by using the generalized inverse of the estimated variance matrix.

In such cases, one can always just drop moment conditions. Of course, then the question is which ones to drop, and in particular whether any rule based on the data will be useful. In work not summarized in this paper, we have investigated the use of principal components, as suggested by Doran and Schmidt (2006) for cases of near-singularity. We were unable to come up with any solid results that would indicate that principal components are useful when the estimated weighting is singular, although this remains a topic worth exploring.

1.6 Appendix: Proof of Result 1.2

Define $G(\theta) = [g(y_1,\theta),...,g(y_n,\theta)]$, $k \times n$. Let e_n be an $n \times 1$ vector of ones. Then $\bar{g}_n(\theta) = \frac{1}{n}G(\theta)e_n$ and $\hat{C}_n(\theta) = \frac{1}{n}G(\theta)G(\theta)'$. Therefore the CUE criterion function is

$$Q_n^{CUE}(\theta) = \frac{1}{n} e_n' G(\theta)' \left[G(\theta) G(\theta)' \right]^+ G(\theta) e_n . \tag{1.9}$$

Now we use the fact that, for A such that AA' is singular but $A'A$ is nonsingular, $(AA')^+ = A(AA')^{-2} A'$. Therefore

$$\begin{aligned} Q_n^{CUE}(\theta) &= \frac{1}{n} e_n' G(\theta)' G(\theta) \left[G(\theta)' G(\theta) \right]^{-2} G(\theta)' G(\theta) e_n \\ &= \frac{1}{n} e_n' e_n \\ &= 1. \end{aligned} \tag{1.10}$$

The proof of part (ii) is exactly the same as for the case of $n = 1$.

1.7 Appendix: Proof of Result 1.3

Let $\hat{\mu} = \bar{\bar{y}}$ be the initial estimate. Let $x_i = (y_i - \hat{\mu} e_k)$ for $i = 1,2$ and $X = [x_1 \quad x_2]$. Let $B = \left[\frac{1}{2} \sum_{i=1}^{2} (y_i - \hat{\mu} e_k)(y_i - \hat{\mu} e_k)' \right]^+ = \left(\frac{1}{2} XX' \right)^+$. Then the two step efficient GMM estimator is

$$\tilde{\mu} = \arg\min_{\mu} (\bar{y} - \hat{\mu} e_k) B (\bar{y} - \hat{\mu} e_k)' = \frac{e_k' B \bar{y}}{e_k' B e_k}.$$

We can rewrite $\tilde{\mu}$ as

$$\begin{aligned} \tilde{\mu} &= \frac{e_k' B \bar{y}}{e_k' B e_k} \\ &= \frac{e_k' B (\frac{1}{2} X e_2 + \hat{\mu} e_k)}{e_k' B e_k} \\ &= \hat{\mu} + \frac{1}{2} \frac{e_k' B X e_2}{e_k' B e_k}. \end{aligned}$$

This implies $\hat{\mu} - \tilde{\mu} = -\frac{1}{2}\frac{e_k' BXe_2}{e_k' Be_k}$ and we can rewrite $(\bar{y} - \tilde{\mu}e_k)$ as $\frac{1}{2}Xe_2 + (\hat{\mu} - \tilde{\mu})e_k$.

Hence,

$$
\begin{aligned}
Q(\tilde{\mu}) &= (\bar{y} - \tilde{\mu}e_k)' B(\bar{y} - \tilde{\mu}e_k) \\
&= \left(\frac{1}{2}Xe_2 + (\hat{\mu} - \tilde{\mu})e_k\right)' B\left(\frac{1}{2}Xe_2 + (\hat{\mu} - \tilde{\mu})e_k\right) \\
&= \frac{1}{4}e_2' X' BXe_2 + (\hat{\mu} - \tilde{\mu})^2 e_k' Be_k + \frac{1}{2}(\hat{\mu} - \tilde{\mu})e_k' BXe_2 + \frac{1}{2}(\hat{\mu} - \tilde{\mu})e_2' X' Be_k \\
&= \frac{1}{4}e_2' X'\left(2X(X'X)^{-2}X'\right)Xe_2 + (\hat{\mu} - \tilde{\mu})^2 e_k' Be_k + (\hat{\mu} - \tilde{\mu})e_2' X' Be_k \\
&= \frac{1}{4}(2e_2' e_2) + \left(-\frac{1}{2}\frac{e_k' BXe_2}{e_k' Be_k}\right)^2 e_k' Be_k + \left(-\frac{1}{2}\frac{e_k' BXe_2}{e_k' Be_k}\right)e_2' X' Be_k \\
&= 1 + \frac{1}{4}\frac{(e_k' BXe_2)^2}{e_k' Be_k} - \frac{1}{2}\frac{(e_k' BXe_2)^2}{e_k' Be_k} \\
&= 1 - \frac{1}{4}\frac{(e_k' BXe_2)^2}{e_k' Be_k} \\
&= 1 - \frac{1}{4}\frac{\left(e_k'\left(2X(X'X)^{-2}X'\right)Xe_2\right)^2}{e_k' Be_k} \\
&= 1 - \frac{1}{2}\frac{\left(e_k' X(X'X)^{-1}e_2\right)^2}{e_k' Be_k} \\
&\le 1.
\end{aligned}
$$

We have equality when

$$e_k' X(X'X)^{-1}e_2 = 0. \tag{1.11}$$

Now we show that this occurs when condition (1.8) of Result 1.3 holds.

We note that

$$e_k X = k\left[(\bar{y}_1 - \hat{\mu}) \quad (\bar{y}_2 - \hat{\mu})\right]. \tag{1.12}$$

Adding and subtracting $\bar{y}_1 e_k$ and $\bar{y}_2 e_k$, and rewriting, we have

$$y_1 - \hat{\mu} = (y_1 - \bar{y}_1 e_k) + (\bar{y}_1 - \hat{\mu}) e_k;$$
$$y_2 - \hat{\mu} = (y_2 - \bar{y}_2 e_k) + (\bar{y}_2 - \hat{\mu}) e_k.$$

Now define $\Delta = \bar{y}_1 - \bar{y}_2$. Then we can rewrite (1.12) as

$$e_k X = \frac{1}{2} k\Delta [1 \quad -1]. \tag{1.13}$$

Lemma 1.1 Define $s_{ij} = (y_i - \bar{y}_i e_k)'(y_j - \bar{y}_j e_k)$ for $i, j = 1,2$. Then

$$X'X = \begin{bmatrix} s_{11} + \frac{1}{4}k\Delta^2 & s_{12} - \frac{1}{4}k\Delta^2 \\ s_{21} - \frac{1}{4}k\Delta^2 & s_{22} + \frac{1}{4}k\Delta^2 \end{bmatrix}.$$

Proof.

$$\begin{aligned}
(y_1 - \hat{y}_1 e_k)'(y_1 - \hat{y}_1 e_k) &= [(y_1 - \bar{y}_1 e_k) + (\bar{y}_1 - \hat{\mu}) e_k]'[(y_1 - \bar{y}_1 e_k) + (\bar{y}_1 - \hat{\mu}) e_k] \\
&= \begin{bmatrix} (y_1 - \bar{y}_1 e_k) + \\ \left(\bar{y}_1 - \frac{1}{2}(\bar{y}_1 - \bar{y}_2)\right) e_k \end{bmatrix}' \begin{bmatrix} (y_1 - \bar{y}_1 e_k) + \\ \left(\bar{y}_1 - \frac{1}{2}(\bar{y}_1 - \bar{y}_2)\right) e_k \end{bmatrix} \\
&= \left[(y_1 - \bar{y}_1 e_k) + \frac{1}{2}\Delta e_k\right]' \left[(y_1 - \bar{y}_1 e_k) + \frac{1}{2}\Delta e_k\right] \\
&= (y_1 - \bar{y}_1 e_k)'(y_1 - \bar{y}_1 e_k) + \Delta e_k'(y_1 - \bar{y}_1 e_k) + \frac{1}{4}k\Delta^2 \\
&= (y_1 - \bar{y}_1 e_k)'(y_1 - \bar{y}_1 e_k) + \frac{1}{4}k\Delta^2 \\
&= s_{11} + \frac{1}{4}k\Delta^2
\end{aligned}$$

similar for the other terms. □

Lemma 1.2

$$\det(X'X) = \det\begin{bmatrix} s_{11} + \frac{1}{4}k\Delta^2 & s_{12} - \frac{1}{4}k\Delta^2 \\ s_{21} - \frac{1}{4}k\Delta^2 & s_{22} + \frac{1}{4}k\Delta^2 \end{bmatrix}$$
$$= \left(s_{11} + \frac{1}{4}k\Delta^2\right)\left(s_{22} + \frac{1}{4}k\Delta^2\right) - \left(s_{12} - \frac{1}{4}k\Delta^2\right)$$
$$= \left(s_{11}s_{22} - s_{12}^2\right) + \frac{1}{4}k\Delta^2\left(s_{11} + s_{22} + 2s_{12}\right)$$

Using Lemma 1.1 and Lemma 1.2, we obtain

$$(X'X)^{-1} = \frac{1}{\det(X'X)}\begin{bmatrix} s_{22} + \frac{1}{4}k\Delta^2 & -s_{12} + \frac{1}{4}k\Delta^2 \\ -s_{21} + \frac{1}{4}k\Delta^2 & s_{11} + \frac{1}{4}k\Delta^2 \end{bmatrix}. \qquad (1.14)$$

Combining (1.12) and (1.14),

$$e_k X(X'X)^{-1} = \frac{1}{2}\frac{k\Delta}{\det(X'X)}\left[s_{22} + s_{12} \quad -s_{12} - s_{11}\right]$$

and then

$$e_k X(X'X)^{-1} e_2 = \frac{1}{2}\frac{k\Delta}{\det(X'X)}(s_{22} - s_{11}). \qquad (1.15)$$

Lemma 1.3 $e_k' B e_k = \frac{1}{2}\frac{(k\Delta)^2}{\left[\det(X'X)\right]^2}\left[(s_{22}+s_{12})^2+(s_{12}+s_{11})^2\right].$

Proof.

$$
\begin{aligned}
e_k' B e_k &= 2e_k' X(X'X)^{-2} X' e_k \\
&= 2\left[e_k' X(X'X)^{-1}\right]\cdot\left[(X'X)^{-1} X' e_k\right] \\
&= 2\left[\frac{1}{2}\frac{k\Delta}{\det(X'X)}\left[s_{22}+s_{12} \quad -s_{12}-s_{11}\right]\right]\left[\frac{1}{2}\frac{k\Delta}{\det(X'X)}\left[s_{22}+s_{12} \quad -s_{12}-s_{11}\right]\right] \\
&= \frac{1}{2}\frac{(k\Delta)^2}{\left[\det(X'X)\right]^2}\left[(s_{22}+s_{12})^2+(s_{12}+s_{11})^2\right]
\end{aligned}
$$

From (1.15) and Lemma 1.3, we have

$$
\frac{\left(e_k' X(X'X)^{-1} e_2\right)^2}{e_k' B e_k} = \frac{(s_{22}-s_{11})^2}{(s_{22}+s_{12})^2+(s_{22}+s_{12})^2}. \tag{1.16}
$$

Therefore, if $s_{11}=s_{22}$, or $\sum_{j=1}^{k}(y_{1j}-\bar{y}_1)=\sum_{j=1}^{k}(y_{2j}-\bar{y}_2)$, then

$$
\frac{\left(e_k' X(X'X)^{-1} e_2\right)^2}{e_k' B e_k} = 0. \qquad \square
$$

Chapter 2

Estimates of Technical Inefficiency in Stochastic Frontier Models with Panel Data: Generalized Panel Jackknife Estimation

2.1 Introduction

In this chapter 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.

In this chapter we consider instead bias corrections based on the jackknife. If the bias of the fixed effects 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 chapter 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 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 "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 chapter is as follows. In Section 2.2, we define some notation and give a brief review of fixed effects estimation of the stochastic frontier model with panel data. In Section 2.3 we show that the bias is of order $T^{-1/2}$ for the case of an exact tie or a "near tie." Section 2.4 describes the generalized panel jackknife that is appropriate in this circumstance. In Section 2.5 we explain the design of our Monte Carlo experiments, and Section 2.6 gives its results. Finally, Section 2.7 contains our concluding remarks.

2.2 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):

$$y_{it} = \alpha + x_{it}'\beta + v_{it} - u_i, i = 1,...,N; t = 1,...,T, \tag{2.1}$$

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 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 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 (2.1) as a standard panel data model:

$$y_{it} = \alpha_i + x_{it}'\beta + v_{it}. \tag{2.2}$$

Obviously, $\alpha_i \leq \alpha$ since $u_i \geq 0$. When α_i (and u_i) is treating as fixed, (2.2) 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}.

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:

$$\hat{\alpha} = \max_{j=1,...,N} \hat{\alpha}_j, \qquad \hat{u}_i = \hat{\alpha} - \hat{\alpha}_i, \qquad i = 1,...,N. \tag{2.3}$$

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 chapter, 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 hope only to 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 $\alpha_{(1)} \leq \alpha_{(2)} \leq ... \leq \alpha_{(N)}$, so that (N) indexes the firm with the largest value of α_i among 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 $u_{(1)} \geq u_{(2)} \geq ... \geq u_{(N)}$. Obviously, $\alpha_{(i)} = \alpha - u_{(i)}$ for all i and specifically $\alpha_{(N)} = \alpha - u_{(N)}$.

Now we define the relative inefficiency measures

$$u_i^* = u_i - u_{(N)} = \alpha_{(N)} - \alpha_i. \tag{2.4}$$

These are the focus of this chapter since, as $T \to \infty$ with N fixed, $\hat{\alpha}$ is a consistent estimate of $\alpha_{(N)}$, not α, and $\hat{u}_i^*$ are consistent estimates 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 $\hat{\alpha} \geq \hat{\alpha}_{(N)}$ and $E(\hat{\alpha}_{(N)}) = \alpha_{(N)}$. That is, the max operator in (2.3) induces 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 the 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

$$\hat{\alpha} - [E(\hat{\alpha}^{boot}) - \hat{\alpha}] = 2\hat{\alpha} - E(\hat{\alpha}^{boot}). \tag{2.5}$$

The feasible version of this is

$$\hat{\alpha}_{BC}^{boot} = 2\hat{\alpha} - B^{-1}\sum_{b=1}^{B}\hat{\alpha}^{(b)}, \tag{2.6}$$

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.

In this chapter, we will consider the jackknife as a simple alternative to the bootstrap.

2.3 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 distinct from all 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:

$$y_{it} = \alpha_i + v_{it}, i = 1,\ldots,N; t = 1,\ldots,T, \tag{2.7}$$

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 $\sqrt{T}(\hat{\alpha}_i - \alpha_i) \to N(0,\sigma^2)$. However, the inclusion of regressors would not alter our results since the within estimator of β is unbiased, and our results really only depend on the vector whose i^{th} element is $\sqrt{T}(\hat{\alpha}_i - \alpha_i)$ being 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.

2.3.1 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 was 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}.

2.3.2 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 $\alpha_{(N)} = \alpha_1 = \alpha_2 = \ldots = \alpha_k$ 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,

$$\begin{aligned}\sqrt{T}(\hat{\alpha} - \alpha_{(N)}) &= \sqrt{T}\max(\hat{\alpha}_1 - \alpha_{(N)}, \hat{\alpha}_2 - \alpha_{(N)}, \ldots, \hat{\alpha}_k - \alpha_{(N)}) \\ &= \max\{\sqrt{T}(\hat{\alpha}_1 - \alpha_{(N)}), \sqrt{T}(\hat{\alpha}_2 - \alpha_{(N)}), \ldots, \sqrt{T}(\hat{\alpha}_k - \alpha_{(N)})\}\end{aligned} \tag{2.8}$$

and therefore $\sqrt{T}(\hat{\alpha} - \alpha_{(N)}) \to Z$ where Z is the maximum of a set of k normals with zero means. 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 expansion for the case of $N = k = 2$ and the simple model above (with no regressors). We first state the following Lemma.

Lemma 2.1 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

$$E[\max(X_1, X_2) - \mu] = (1/\sqrt{\pi})\sigma. \tag{2.9}$$

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, $\rho = \sigma^2 \big/ \sqrt{2\sigma^2\sigma^2} = 1/\sqrt{2}$ 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 $T^{-1/2}$.

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

$$\begin{aligned}\sqrt{T}[\max(\hat{\alpha}_1,\hat{\alpha}_2) - \alpha_1] &= \max[\sqrt{T}(\hat{\alpha}_1 - \alpha_1), \sqrt{T}(\hat{\alpha}_2 - \alpha_1)] \\ &\to Z,\end{aligned} \tag{2.10}$$

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).

Lemma 2.2 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

$$E[\max(X_1,X_2)] = [\Phi(\mu_*/\sqrt{2})\mu_* + \sqrt{2}\phi(\mu_*/\sqrt{2})]\sigma, \tag{2.11}$$

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

Proof. See Appendix.

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

$$bias = [\Phi(-c/\sqrt{2}\sigma)(-c/\sigma) + \sqrt{2}\phi(-c/\sqrt{2}\sigma) + (c/\sigma)]T^{-1/2}\sigma, \tag{2.12}$$

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

2.4 Correcting Bias with the Panel Jackknife and the Generalized Panel Jackknife

2.4.1 The Panel Jackknife

Jackknife estimation is an automatic bias reduction tool under assumption 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 comprehensive background on jackknife estimation, see Miller (1974).

To describe the jackknife in a general setting, let the data be indexed by $t = 1,2,\ldots,T$. 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

$$J(\hat{\theta}) = T\hat{\theta} - (T-1)T^{-1}\sum_t \hat{\theta}_{(t)}. \tag{2.13}$$

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

$$E(\hat{\theta}) = \theta + T^{-1}B + T^{-2}D + O(T^{-3}). \tag{2.14}$$

Then

$$E[J(\hat{\theta})] = \theta + \left(\frac{1}{T} - \frac{1}{T-1}\right)D + O(T^{-2}) = \theta + O(T^{-2}). \tag{2.15}$$

So if the bias is of order T^{-1}, in the sense that (2.14) 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 nonlinear dynamic panel data models. In the panel data setting, even though there are really NT observations, we treat the number of observations in (2.13) 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 split-sample jackknife estimator:

$$SSJ(\hat{\theta}) = 2\hat{\theta} - (1/2)(\hat{\theta}^{(1)} + \hat{\theta}^{(2)}), \tag{2.16}$$

where $\hat{\theta}$ is the fixed effects estimator based on the full sample; $\hat{\theta}^{(1)}$ and $\hat{\theta}^{(2)}$ based on first- and second- half of panel sample, where each half-panel consists of $T/2$ consecutive observations over time for all cross-sectional units. They show that the split-sample jackknife estimator, which is an extension of the panel jackknife with $T = 2$, also removes the bias of order T^{-1} from the fixed effects 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.

2.4.2 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 the 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

$$G(\hat{\theta}_1, \hat{\theta}_2) = \frac{\hat{\theta}_1 - R\hat{\theta}_2}{1 - R}. \tag{2.17}$$

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 2.1 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 leading term of known order. Then the generalized jackknife removes the first term of the bias expansion.

<u>**Theorem 2.2**</u> 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 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 proof of Theorem 2.1. □

2.4.3 The Generalized Panel Jackknife When 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:

$$E(\hat{\theta}) = \theta + T^{-1/2}B + T^{-1}D + O(T^{-3/2}). \tag{2.18}$$

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 2.2 is equal to

$$R = (B/\sqrt{T})/(B/\sqrt{T-1}) = \sqrt{T-1}/\sqrt{T} \tag{2.19}$$

and the generalized jackknife is

$$G(\hat{\theta}) = \frac{\sqrt{T}}{\sqrt{T}-\sqrt{T-1}}\hat{\theta} - \frac{\sqrt{T-1}}{\sqrt{T}-\sqrt{T-1}}T^{-1}\sum_t \hat{\theta}_{(t)}. \tag{2.20}$$

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

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

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

2.4.4 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 2.3 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

$$E[J(\hat{\theta})] = \theta + B\sqrt{T} - B\sqrt{T-1} = \theta + \frac{B}{\sqrt{T} + \sqrt{T-1}}. \tag{2.21}$$

Comparing the bias on (2.21) to the original bias of $T^{-1/2}B$, we have removed about half of the first-order bias term. □

Theorem 2.4 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 $E(\hat{\theta}) = \theta + T^{-1/2}B + \textit{higher order terms}$. So, again dropping the higher order terms,

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

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

$$= \theta + \frac{1}{\sqrt{T}-\sqrt{T-1}}[T^{-1/2} - (T-1)^{-1/2}]B \qquad (2.22)$$

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

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

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

2.5 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 given in (2.7). Thus, the data generating process is

$$\begin{aligned} y_{it} &= \alpha + v_{it} - u_i \\ &= \alpha_i + v_{it}, i = 1,\ldots,N; t = 1,\ldots,T, \end{aligned} \tag{2.23}$$

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 MSEs 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 $(\sigma_u^2)_* = \text{var}(u) = ((\pi - 2)/\pi)\sigma_u^2$. We will use instead the parameter μ_* defined by

$$\mu_* = \frac{(\sigma_u)_*}{T^{-1/2}\sigma_v}. \tag{2.24}$$

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 $\hat{\alpha}_i$. So our parameter space is $\{\mu_*, N, T\}$.

We set scale by setting $\sigma_v^2/T = 0.1$. 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 can be determined:

(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 shown in Supplemental Table 2.13 - 2.48.

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 panel jackknife or the bootstrap to be better than the generalized panel 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

$$\alpha_{(N)} - T^{-1/2}[\alpha_{(N)} - \alpha_{(N-1)}]. \tag{2.25}$$

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 point 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) $\hat{\theta}_r =$ estimate of θ_r 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:

$$bias(\hat{\theta}) = NREP^{-1}\sum_r (\hat{\theta}_r - \theta_r) = \bar{\hat{\theta}} - \hat{\theta}. \tag{2.26}$$

Then we define the mean squared error as

$$\begin{aligned} MSE(\hat{\theta}) &= NREP^{-1}\sum_r (\hat{\theta}_r - \theta_r)^2 \\ &= NREP^{-1}\sum_r [(\hat{\theta}_r - \theta_r) - bias(\hat{\theta})]^2 + bias(\hat{\theta})^2 \end{aligned} \tag{2.27}$$

and the variance as

$$\operatorname{var}(\hat{\theta}) = MSE(\hat{\theta}) - bias(\hat{\theta})^2$$
$$= NREP^{-1}\sum_r[(\hat{\theta}_r - \theta_r) - bias(\hat{\theta})]^2. \tag{2.28}$$

2.6 Results of the Monte Carlo Experiments

Tables 2.1, 2.2, and 2.3 give the results of Experiment I in which there is no tie. All of these results are for $T = 10$. Table 2.1 gives the bias of the estimates, while Table 2.2 gives variance and Table 2.3 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 2.1, 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 "no 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 2.3. 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 2.4.

Table 2.2 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 2.3 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 2.4, 2.5, and 2.6.

In terms of bias, we see in Table 2.4 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 2.3, 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 2.5, the variances of the estimates are rather similar to the variances for the case of no tie (Table 2.2). 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 2.2 ($\hat{\alpha}$, bias-corrected bootstrap, panel jackknife and generalized panel jackknife).

In terms of MSE, we see in Table 2.6 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 2.7, 2.8, and 2.9. 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 2.7 are quite similar to those of Table 2.4 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 2.7 that the generalized panel jackknife has the smallest bias.

In terms of variance (Table 2.8) and MSE (Table 2.9), 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 Supplemental Table 2.13 – 2.48. In this chapter we will display the results only for $\mu_* = 1$ and $N = 20$. Tables 2.10, 2.11, and 2.12 give the bias, variance, and MSE of the various estimates.

As discussed in Section 2.5, 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 2.10 – 2.12 is not surprising. In Table 2.10, the effect of changing T on the bias of the estimates is very minor. In Table 2.11, 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 2.12 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.

2.7 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 chapter 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.

2.8 Output Tables

Table 2.1: (Experiment I: No 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	0.1671	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 2.2: (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 2.3: (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 2.4: (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 2.5: (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 2.6: (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 2.7: (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 2.8: (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 2.9: (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 2.10: (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 2.11: (Effect of Chaning 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	5	0.0365	0.1208	0.3162	0.0647
NO	10	0.0359	0.1534	0.4191	0.0650
	20	0.0365	0.1678	0.4717	0.0669
TIE	50	0.0363	0.2247	0.7607	0.0663
	100	0.0378	0.3265	1.0631	0.0702
II	5	0.0334	0.1139	0.3029	0.0593
EXACT	10	0.0351	0.1484	0.4048	0.0640
	20	0.0361	0.1867	0.5337	0.0650
TIE	50	0.0384	0.2568	0.7893	0.0714
	100	0.0366	0.3046	0.9902	0.0680
III	5	0.0332	0.1137	0.3030	0.0590
NEAR	10	0.0351	0.1483	0.4050	0.0639
	20	0.0360	0.1868	0.5343	0.0649
TIE	50	0.0384	0.2542	0.7810	0.0713
	100	0.0366	0.3089	1.0055	0.0680

Table 2.12: (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

2.9 Appendix: Proof of Lemma 2.2

Lemma 2.3 Let X_1 and X_2 be independent bivariate normals with different mean, but identical variance, i.e.,

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left[\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}\right]. \tag{2.29}$$

Then

$$E[\max(X_1,X_2)] = P(X_1 \geq X_2)\cdot E(X_1 \mid X_1 > X_2) + P(X_2 \geq X_1)\cdot E(X_2 \mid X_2 > X_1).$$

Proof. Let $m = \max(X_1, X_2)$, $s = 1$ if $X_1 \geq X_2$ and $= 2$ if $X_2 \geq X_1$, and $m = X_1$ if $s = 1$ and $= X_2$ if $s = 2$. Then

$$\begin{aligned} f(m) &= f(m, s=1) + f(m, s=2) \\ &= P(s=1) f(m \mid s=1) + P(s=2) f(m \mid s=2) \end{aligned}$$

and

$$\begin{aligned} E(m) &= \int m f(m) dm \\ &= P(s=1)E(m \mid s=1) + P(s=2)E(m \mid s=2) \\ &= P(s=1)E(X_1 \mid X_1 \geq X_2) + P(s=2)E(X_2 \mid X_2 \geq X_1) \\ &= P(X_1 \geq X_2)E(X_1 \mid X_1 \geq X_2) + P(X_2 \geq X_1)E(X_2 \mid X_2 \geq X_1). \end{aligned}$$

□

Lemma 2.4 By Lemma 2.3,

$$\begin{aligned} E[\max(X_1,X_2)] = {} & \Phi\left[\frac{(\mu_1-\mu_2)}{\sqrt{2\sigma^2}}\right]\cdot\left[\mu_1 + \frac{1}{\sqrt{2}}\sigma\lambda\left(-\frac{(\mu_1-\mu_2)}{\sqrt{2\sigma^2}}\right)\right] \\ & + \Phi\left[\frac{(\mu_2-\mu_1)}{\sqrt{2\sigma^2}}\right]\cdot\left[\mu_2 + \frac{1}{\sqrt{2}}\sigma\lambda\left(-\frac{(\mu_2-\mu_1)}{\sqrt{2\sigma^2}}\right)\right]. \end{aligned}$$

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

$$E[\max(X_1,X_2)]=\left[\Phi\left(\frac{\mu_*}{\sqrt{2}}\right)\mu_*+\sqrt{2}\phi\left(\frac{\mu_*}{\sqrt{2}}\right)\right]\sigma, \tag{2.30}$$

where $\mu_*=\dfrac{\mu_2}{\sigma}$.

Proof. By Lemma 2.3, Without loss of generality, let $\mu_1=0$ and μ_* be some constant such that $\mu_2=\mu_*\sigma$ (since "σ" is $\dfrac{\sigma}{\sqrt{T}}$, $\mu_2-\mu_1\to 0$ as $T\to\infty$). Then

$$\begin{aligned}
P(X_1\ge X_2)&=P(X_1-X_2\ge 0)\\
&=P\left[\frac{(X_1-X_2)-(\mu_1-\mu_2)}{\sqrt{2\sigma^2}}\ge\frac{0-(\mu_1-\mu_2)}{\sqrt{2\sigma^2}}\right]\\
&=1-\Phi\left[-\frac{(\mu_1-\mu_2)}{\sqrt{2\sigma^2}}\right]\\
&=\Phi\left[\frac{(\mu_1-\mu_2)}{\sqrt{2\sigma^2}}\right].
\end{aligned}$$

Similarly,

$$P(X_2\ge X_1)=\Phi\left[\frac{(\mu_2-\mu_1)}{\sqrt{2\sigma^2}}\right].$$

From the facts about truncated bivariate normal distribution,

$$E(X_1\mid X_1>X_2)=\mu_1+\frac{1}{\sqrt{2}}\sigma\lambda\left(-\frac{(\mu_1-\mu_2)}{\sqrt{2\sigma^2}}\right),$$

$$E(X_2\mid X_2>X_1)=\mu_2+\frac{1}{\sqrt{2}}\sigma\lambda\left(-\frac{(\mu_2-\mu_1)}{\sqrt{2\sigma^2}}\right).$$

Therefore,

$$E[\max(X_1,X_2)] = \Phi\left[\frac{(\mu_1-\mu_2)}{\sqrt{2\sigma^2}}\right]\cdot\left[\mu_1 + \frac{1}{\sqrt{2}}\sigma\lambda\left(-\frac{(\mu_1-\mu_2)}{\sqrt{2\sigma^2}}\right)\right]$$
$$+\Phi\left[\frac{(\mu_2-\mu_1)}{\sqrt{2\sigma^2}}\right]\cdot\left[\mu_2 + \frac{1}{\sqrt{2}}\sigma\lambda\left(-\frac{(\mu_2-\mu_1)}{\sqrt{2\sigma^2}}\right)\right].$$

Since we assume $\mu_1 = 0$ and $\mu_2 = \mu_* \sigma$ (hence, $\mu_* = \frac{\mu_2}{\sigma}$), $\frac{(\mu_2-\mu_1)}{\sqrt{2\sigma^2}} = -\frac{\mu_*}{\sqrt{2}}$;

$\lambda\left(\frac{\mu_*}{\sqrt{2}}\right) = \frac{\varphi\left(\frac{\mu_*}{\sqrt{2}}\right)}{\Phi\left(-\frac{\mu_*}{\sqrt{2}}\right)}$; and $\lambda\left(-\frac{\mu_*}{\sqrt{2}}\right) = \frac{\varphi\left(\frac{\mu_*}{\sqrt{2}}\right)}{\Phi\left(\frac{\mu_*}{\sqrt{2}}\right)}$. Then,

$$\begin{aligned}
E[\max(X_1,X_2)] &= \Phi\left(-\frac{\mu_*}{\sqrt{2}}\right)\left[\frac{1}{\sqrt{2}}\sigma\lambda\left(\frac{\mu_*}{\sqrt{2}}\right)\right] + \Phi\left(\frac{\mu_*}{\sqrt{2}}\right)\left[\mu_*\sigma + \frac{1}{\sqrt{2}}\sigma\lambda\left(-\frac{\mu_*}{\sqrt{2}}\right)\right] \\
&= \Phi\left(-\frac{\mu_*}{\sqrt{2}}\right)\left[\frac{1}{\sqrt{2}}\sigma\frac{\phi\left(\frac{\mu_*}{\sqrt{2}}\right)}{\Phi\left(-\frac{\mu_*}{\sqrt{2}}\right)}\right] + \Phi\left(\frac{\mu_*}{\sqrt{2}}\right)\left[\mu_*\sigma + \frac{1}{\sqrt{2}}\sigma\frac{\phi\left(\frac{\mu_*}{\sqrt{2}}\right)}{\Phi\left(\frac{\mu_*}{\sqrt{2}}\right)}\right] \\
&= \frac{1}{\sqrt{2}}\sigma\phi\left(\frac{\mu_*}{\sqrt{2}}\right) + \Phi\left(\frac{\mu_*}{\sqrt{2}}\right)\mu_*\sigma + \frac{1}{\sqrt{2}}\sigma\phi\left(\frac{\mu_*}{\sqrt{2}}\right) \\
&= 2\cdot\frac{1}{\sqrt{2}}\sigma\phi\left(\frac{\mu_*}{\sqrt{2}}\right) + \Phi\left(\frac{\mu_*}{\sqrt{2}}\right)\mu_*\sigma \\
&= \left[\Phi\left(\frac{\mu_*}{\sqrt{2}}\right)\mu_* + \sqrt{2}\phi\left(\frac{\mu_*}{\sqrt{2}}\right)\right]\sigma.
\end{aligned}$$

□

Lemma 2.5 By Lemma 2.4, the bias is proportional to σ when means are unequal:

$$bias = \left[\Phi\left(\frac{\mu_*}{\sqrt{2}}\right)\mu_* + \sqrt{2}\phi\left(\frac{\mu_*}{\sqrt{2}}\right) - \mu_*\right]\cdot\sigma. \tag{2.31}$$

Proof.

$$\begin{aligned}
bias &= E[\max(X_1, X_2)] - \mu_2 \\
&= E[\max(X_1, X_2)] - \mu_*\sigma \\
&= \left[\Phi\left(\frac{\mu_*}{\sqrt{2}}\right)\mu_* + \sqrt{2}\phi\left(\frac{\mu_*}{\sqrt{2}}\right)\right]\cdot\sigma - \mu_*\sigma \\
&= \left[\Phi\left(\frac{\mu_*}{\sqrt{2}}\right)\mu_* + \sqrt{2}\phi\left(\frac{\mu_*}{\sqrt{2}}\right) - \mu_*\right]\cdot\sigma
\end{aligned}$$

□

2.10 Supplementary Tables

Supplemental Table 2.13: (Experiment I: No Tie) $T = 5$, 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	0.1597	0.0737	-0.0225	0.0996
$10^{-1/2}$	2	0.1317	0.0513	-0.0386	0.0747
1	2	0.0829	0.0251	-0.0395	0.0397
$10^{1/2}$	2	0.0384	0.0060	-0.0301	0.0157
10	2	0.0185	0.0059	-0.0081	0.0099
10^{-1}	10	0.4632	0.2342	-0.0219	0.2991
$10^{-1/2}$	10	0.4028	0.1771	-0.0753	0.2420
1	10	0.2868	0.0907	-0.1285	0.1455
$10^{1/2}$	10	0.1524	0.0220	-0.1238	0.0564
10	10	0.0623	0.0006	-0.0684	0.0167
10^{-1}	20	0.5551	0.2601	-0.0697	0.3519
$10^{-1/2}$	20	0.4961	0.2125	-0.1046	0.2988
1	20	0.3842	0.1333	-0.1472	0.2098
$10^{1/2}$	20	0.2477	0.0785	-0.1107	0.1219
10	20	0.1230	0.0273	-0.0797	0.0496
10^{-1}	50	0.6730	0.3228	-0.0687	0.4283
$10^{-1/2}$	50	0.6119	0.2734	-0.1050	0.3727
1	50	0.5012	0.1999	-0.1371	0.2855
$10^{1/2}$	50	0.3627	0.1319	-0.1262	0.1934
10	50	0.2179	0.0730	-0.0890	0.1044
10^{-1}	100	0.7429	0.3411	-0.1081	0.4638
$10^{-1/2}$	100	0.6827	0.2915	-0.1459	0.4099
1	100	0.5769	0.2311	-0.1556	0.3288
$10^{1/2}$	100	0.4343	0.1492	-0.1694	0.2263
10	100	0.2853	0.0786	-0.1525	0.1307

Supplemental Table 2.14: (Experiment I: No Tie) $T = 5$, 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	0.0676	0.1030	0.1813	0.0804
$10^{-1/2}$	2	0.0704	0.1063	0.1840	0.0836
1	2	0.0790	0.1088	0.1732	0.0909
$10^{1/2}$	2	0.0886	0.1056	0.1445	0.0953
10	2	0.0939	0.1033	0.1236	0.0970
10^{-1}	10	0.0351	0.1090	0.2776	0.0603
$10^{-1/2}$	10	0.0373	0.1153	0.2929	0.0636
1	10	0.0462	0.1182	0.2769	0.0731
$10^{1/2}$	10	0.0629	0.1204	0.2468	0.0841
10	10	0.0772	0.1054	0.1712	0.0881
10^{-1}	20	0.0281	0.1202	0.3363	0.0551
$10^{-1/2}$	20	0.0295	0.1207	0.3347	0.0568
1	20	0.0365	0.1208	0.3162	0.0647
$10^{1/2}$	20	0.0537	0.1194	0.2674	0.0789
10	20	0.0767	0.1210	0.2230	0.0942
10^{-1}	50	0.0207	0.1088	0.3244	0.0447
$10^{-1/2}$	50	0.0225	0.1128	0.3314	0.0479
1	50	0.0280	0.1207	0.3420	0.0560
$10^{1/2}$	50	0.0383	0.1203	0.3082	0.0662
10	50	0.0570	0.1187	0.2561	0.0815
10^{-1}	100	0.0180	0.1113	0.3406	0.0421
$10^{-1/2}$	100	0.0194	0.1146	0.3466	0.0451
1	100	0.0224	0.1144	0.3351	0.0488
$10^{1/2}$	100	0.0298	0.1204	0.3306	0.0567
10	100	0.0439	0.1187	0.2867	0.0694

Supplemental Table 2.15: (Experiment I: No Tie) $T = 5$, 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	0.0931	0.1084	0.1818	0.0903
$10^{-1/2}$	2	0.0877	0.1089	0.1855	0.0892
1	2	0.0859	0.1094	0.1748	0.0925
$10^{1/2}$	2	0.0901	0.1056	0.1454	0.0956
10	2	0.0943	0.1033	0.1237	0.0971
10^{-1}	10	0.2497	0.1638	0.2781	0.1498
$10^{-1/2}$	10	0.1996	0.1467	0.2986	0.1222
1	10	0.1285	0.1265	0.2934	0.0943
$10^{1/2}$	10	0.0862	0.1209	0.2621	0.0873
10	10	0.0811	0.1054	0.1759	0.0884
10^{-1}	20	0.3362	0.1878	0.3411	0.1789
$10^{-1/2}$	20	0.2756	0.1658	0.3457	0.1461
1	20	0.1841	0.1386	0.3379	0.1087
$10^{1/2}$	20	0.1150	0.1256	0.2797	0.0937
10	20	0.0919	0.1217	0.2294	0.0967
10^{-1}	50	0.4736	0.2130	0.3291	0.2282
$10^{-1/2}$	50	0.3969	0.1875	0.3424	0.1868
1	50	0.2792	0.1606	0.3608	0.1375
$10^{1/2}$	50	0.1698	0.1377	0.3241	0.1036
10	50	0.1045	0.1241	0.2640	0.0923
10^{-1}	100	0.5699	0.2276	0.3523	0.2573
$10^{-1/2}$	100	0.4855	0.1996	0.3679	0.2131
1	100	0.3552	0.1678	0.3594	0.1569
$10^{1/2}$	100	0.2184	0.1427	0.3594	0.1079
10	100	0.1253	0.1249	0.3100	0.0865

Supplemental Table 2.16: (Experiment II: Exact Tie) $T = 5$, 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	0.1760	0.0892	-0.0078	0.1154
10^{-1}	10	0.4669	0.2387	-0.0164	0.3027
$10^{-1/2}$	10	0.4159	0.1941	-0.0538	0.2554
1	10	0.3292	0.1382	-0.0753	0.1884
$10^{1/2}$	10	0.2506	0.1059	-0.0560	0.1454
10	10	0.2138	0.1097	-0.0067	0.1366
10^{-1}	20	0.5569	0.2641	-0.0633	0.3540
$10^{-1/2}$	20	0.5014	0.2194	-0.0959	0.3039
1	20	0.4032	0.1566	-0.1191	0.2272
$10^{1/2}$	20	0.3005	0.1210	-0.0797	0.1671
10	20	0.2372	0.1130	-0.0259	0.1449
10^{-1}	50	0.6713	0.3183	-0.0765	0.4249
$10^{-1/2}$	50	0.6100	0.2647	-0.1213	0.3678
1	50	0.5049	0.2005	-0.1400	0.2867
$10^{1/2}$	50	0.3836	0.1460	-0.1196	0.2090
10	50	0.2853	0.1159	-0.0734	0.1594
10^{-1}	100	0.7427	0.3407	-0.1088	0.4635
$10^{-1/2}$	100	0.6841	0.2961	-0.1377	0.4118
1	100	0.5831	0.2339	-0.1565	0.3352
$10^{1/2}$	100	0.4643	0.1839	-0.1297	0.2605
10	100	0.3489	0.1456	-0.0817	0.1955

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

Supplemental Table 2.17: (Experiment II: Exact Tie) $T = 5$, Variance of the Estimates

μ_*	N	(1) $\hat{\alpha}$	(2) $J(\hat{\alpha})$	(3) $G(\hat{\alpha})$	(4) $\hat{\alpha}_{BC}^{boot}$
**	2	0.0670	0.1014	0.1776	0.0797
10^{-1}	10	0.0350	0.1104	0.2845	0.0601
$10^{-1/2}$	10	0.0369	0.1132	0.2893	0.0618
1	10	0.0457	0.1197	0.2878	0.0718
$10^{1/2}$	10	0.0591	0.1234	0.2657	0.0825
10	10	0.0661	0.1106	0.2060	0.0837
10^{-1}	20	0.0278	0.1197	0.3366	0.0544
$10^{-1/2}$	20	0.0286	0.1162	0.3238	0.0548
1	20	0.0334	0.1139	0.3029	0.0593
$10^{1/2}$	20	0.0446	0.1131	0.2670	0.0689
10	20	0.0548	0.1070	0.2226	0.0738
10^{-1}	50	0.0205	0.1099	0.3314	0.0444
$10^{-1/2}$	50	0.0216	0.1147	0.3442	0.0465
1	50	0.0250	0.1129	0.3246	0.0504
$10^{1/2}$	50	0.0330	0.1093	0.2901	0.0578
10	50	0.0469	0.1148	0.2662	0.0708
10^{-1}	100	0.0179	0.1116	0.3434	0.0418
$10^{-1/2}$	100	0.0186	0.1095	0.3335	0.0426
1	100	0.0233	0.1187	0.3484	0.0500
$10^{1/2}$	100	0.0312	0.1210	0.3304	0.0597
10	100	0.0414	0.1149	0.2834	0.0670

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

Supplemental Table 2.18: (Experiment II: Exact Tie) $T = 5$, MSE of the Estimates

μ_*	N	(1) $\hat{\alpha}$	(2) $J(\hat{\alpha})$	(3) $G(\hat{\alpha})$	(4) $\hat{\alpha}_{BC}^{boot}$
**	2	0.0980	0.1094	0.1777	0.0930
10^{-1}	10	0.2530	0.1674	0.2848	0.1517
$10^{-1/2}$	10	0.2098	0.1509	0.2922	0.1270
1	10	0.1541	0.1388	0.2935	0.1073
$10^{1/2}$	10	0.1219	0.1347	0.2689	0.1036
10	10	0.1118	0.1226	0.2061	0.1023
10^{-1}	20	0.3379	0.1895	0.3406	0.1797
$10^{-1/2}$	20	0.2799	0.1644	0.3330	0.1471
1	20	0.1959	0.1384	0.3171	0.1109
$10^{1/2}$	20	0.1349	0.1278	0.2734	0.0968
10	20	0.1110	0.1198	0.2232	0.0947
10^{-1}	50	0.4712	0.2112	0.3373	0.2249
$10^{-1/2}$	50	0.3936	0.1847	0.3590	0.1817
1	50	0.2799	0.1531	0.3441	0.1326
$10^{1/2}$	50	0.1802	0.1306	0.3044	0.1015
10	50	0.1283	0.1282	0.2716	0.0962
10^{-1}	100	0.5695	0.2276	0.3552	0.2566
$10^{-1/2}$	100	0.4866	0.1972	0.3524	0.2121
1	100	0.3634	0.1734	0.3729	0.1624
$10^{1/2}$	100	0.2468	0.1548	0.3472	0.1275
10	100	0.1631	0.1361	0.2901	0.1052

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

Supplemental Table 2.19: (Experiment III: Near Tie) $T = 5$, 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	0.1689	0.0835	-0.0121	0.1088
$10^{-1/2}$	2	0.1547	0.0712	-0.0221	0.0955
1	2	0.1193	0.0443	-0.0396	0.0650
$10^{1/2}$	2	0.0652	0.0136	-0.0441	0.0272
10	2	0.0262	-0.0006	-0.0305	0.0084
10^{-1}	10	0.4666	0.2383	-0.0169	0.3023
$10^{-1/2}$	10	0.4146	0.1925	-0.0558	0.2540
1	10	0.3229	0.1318	-0.0817	0.1816
$10^{1/2}$	10	0.2206	0.0744	-0.0890	0.1144
10	10	0.1319	0.0439	-0.0545	0.0634
10^{-1}	20	0.5568	0.2640	-0.0634	0.3540
$10^{-1/2}$	20	0.5011	0.2192	-0.0960	0.3036
1	20	0.4013	0.1547	-0.1210	0.2252
$10^{1/2}$	20	0.2887	0.1071	-0.0960	0.1540
10	20	0.1912	0.0664	-0.0732	0.0988
10^{-1}	50	0.6713	0.3183	-0.0764	0.4249
$10^{-1/2}$	50	0.6099	0.2646	-0.1214	0.3677
1	50	0.5045	0.2000	-0.1404	0.2863
$10^{1/2}$	50	0.3807	0.1425	-0.1239	0.2056
10	50	0.2694	0.0980	-0.0936	0.1422
10^{-1}	100	0.7427	0.3407	-0.1088	0.4635
$10^{-1/2}$	100	0.6840	0.2960	-0.1378	0.4117
1	100	0.5830	0.2337	-0.1569	0.3350
$10^{1/2}$	100	0.4632	0.1819	-0.1325	0.2590
10	100	0.3422	0.1365	-0.0934	0.1877

Supplemental Table 2.20: (Experiment III: Near Tie) $T = 5$, 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	0.0670	0.1008	0.1758	0.0796
$10^{-1/2}$	2	0.0675	0.1013	0.1763	0.0800
1	2	0.0716	0.1068	0.1816	0.0849
$10^{1/2}$	2	0.0821	0.1087	0.1664	0.0930
10	2	0.0909	0.1062	0.1405	0.0965
10^{-1}	10	0.0350	0.1104	0.2845	0.0601
$10^{-1/2}$	10	0.0368	0.1132	0.2894	0.0618
1	10	0.0457	0.1184	0.2839	0.0716
$10^{1/2}$	10	0.0623	0.1262	0.2694	0.0857
10	10	0.0781	0.1132	0.1946	0.0918
10^{-1}	20	0.0278	0.1197	0.3366	0.0544
$10^{-1/2}$	20	0.0285	0.1162	0.3235	0.0547
1	20	0.0332	0.1137	0.3030	0.0590
$10^{1/2}$	20	0.0448	0.1139	0.2699	0.0690
10	20	0.0602	0.1135	0.2321	0.0794
10^{-1}	50	0.0205	0.1099	0.3314	0.0444
$10^{-1/2}$	50	0.0216	0.1147	0.3443	0.0465
1	50	0.0250	0.1129	0.3244	0.0504
$10^{1/2}$	50	0.0330	0.1088	0.2890	0.0577
10	50	0.0478	0.1158	0.2680	0.0715
10^{-1}	100	0.0179	0.1116	0.3433	0.0418
$10^{-1/2}$	100	0.0186	0.1095	0.3334	0.0426
1	100	0.0233	0.1186	0.3482	0.0500
$10^{1/2}$	100	0.0311	0.1207	0.3295	0.0595
10	100	0.0416	0.1156	0.2856	0.0672

Supplemental Table 2.21: (Experiment III: Near Tie) $T = 5$, 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	0.0956	0.1078	0.1759	0.0914
$10^{-1/2}$	2	0.0914	0.1063	0.1768	0.0892
1	2	0.0859	0.1088	0.1831	0.0891
$10^{1/2}$	2	0.0864	0.1089	0.1684	0.0937
10	2	0.0916	0.1062	0.1414	0.0966
10^{-1}	10	0.2527	0.1672	0.2848	0.1515
$10^{-1/2}$	10	0.2087	0.1503	0.2925	0.1263
1	10	0.1499	0.1358	0.2906	0.1046
$10^{1/2}$	10	0.1109	0.1318	0.2773	0.0988
10	10	0.0954	0.1151	0.1976	0.0958
10^{-1}	20	0.3378	0.1895	0.3406	0.1797
$10^{-1/2}$	20	0.2796	0.1642	0.3327	0.1469
1	20	0.1942	0.1376	0.3176	0.1097
$10^{1/2}$	20	0.1281	0.1253	0.2791	0.0927
10	20	0.0968	0.1179	0.2374	0.0892
10^{-1}	50	0.4711	0.2112	0.3372	0.2249
$10^{-1/2}$	50	0.3936	0.1847	0.3590	0.1817
1	50	0.2795	0.1529	0.3442	0.1324
$10^{1/2}$	50	0.1779	0.1291	0.3043	0.1000
10	50	0.1204	0.1254	0.2767	0.0918
10^{-1}	100	0.5695	0.2276	0.3552	0.2566
$10^{-1/2}$	100	0.4865	0.1971	0.3524	0.2121
1	100	0.3632	0.1732	0.3728	0.1622
$10^{1/2}$	100	0.2457	0.1538	0.3470	0.1266
10	100	0.1587	0.1343	0.2943	0.1024

Supplemental Table 2.22: (Experiment I: No Tie) $T = 20$, 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	0.1788	0.0873	-0.0065	0.1083
$10^{-1/2}$	2	0.1503	0.0686	-0.0151	0.0830
1	2	0.0942	0.0249	-0.0463	0.0396
$10^{1/2}$	2	0.0425	0.0089	-0.0255	0.0131
10	2	0.0258	0.0170	0.0079	0.0169
10^{-1}	10	0.4581	0.2121	-0.0403	0.2643
$10^{-1/2}$	10	0.3993	0.1589	-0.0877	0.2099
1	10	0.2920	0.0953	-0.1064	0.1291
$10^{1/2}$	10	0.1630	0.0354	-0.0956	0.0524
10	10	0.0675	0.0145	-0.0398	0.0149
10^{-1}	20	0.5581	0.2761	-0.0132	0.3253
$10^{-1/2}$	20	0.4964	0.2231	-0.0573	0.2695
1	20	0.3795	0.1417	-0.1023	0.1782
$10^{1/2}$	20	0.2370	0.0607	-0.1203	0.0897
10	20	0.1213	0.0333	-0.0569	0.0391
10^{-1}	50	0.6849	0.3532	0.0130	0.4082
$10^{-1/2}$	50	0.6243	0.3127	-0.0070	0.3551
1	50	0.5097	0.2218	-0.0736	0.2641
$10^{1/2}$	50	0.3668	0.1354	-0.1020	0.1696
10	50	0.2199	0.0614	-0.1011	0.0818
10^{-1}	100	0.7621	0.3897	0.0076	0.4544
$10^{-1/2}$	100	0.6999	0.3367	-0.0360	0.3979
1	100	0.5852	0.2448	-0.1045	0.3040
$10^{1/2}$	100	0.4414	0.1450	-0.1592	0.2038
10	100	0.2923	0.0805	-0.1367	0.1133

Supplemental Table 2.23: (Experiment I: No Tie) $T = 20$, 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	0.0721	0.1343	0.2703	0.0876
$10^{-1/2}$	2	0.0735	0.1314	0.2577	0.0885
1	2	0.0824	0.1396	0.2634	0.0971
$10^{1/2}$	2	0.0971	0.1242	0.1806	0.1069
10	2	0.1014	0.1069	0.1203	0.1034
10^{-1}	10	0.0347	0.1811	0.5272	0.0610
$10^{-1/2}$	10	0.0381	0.1836	0.5215	0.0663
1	10	0.0484	0.1768	0.4621	0.0782
$10^{1/2}$	10	0.0679	0.1619	0.3663	0.0934
10	10	0.0853	0.1231	0.2109	0.0995
10^{-1}	20	0.0287	0.1893	0.5668	0.0577
$10^{-1/2}$	20	0.0301	0.1886	0.5601	0.0598
1	20	0.0365	0.1678	0.4717	0.0669
$10^{1/2}$	20	0.0516	0.1688	0.4344	0.0804
10	20	0.0754	0.1396	0.2853	0.0950
10^{-1}	50	0.0221	0.1916	0.5987	0.0504
$10^{-1/2}$	50	0.0225	0.1879	0.5863	0.0503
1	50	0.0271	0.1906	0.5753	0.0566
$10^{1/2}$	50	0.0405	0.1829	0.5033	0.0733
10	50	0.0630	0.1773	0.4247	0.0937
10^{-1}	100	0.0187	0.2177	0.7015	0.0474
$10^{-1/2}$	100	0.0189	0.2004	0.6434	0.0471
1	100	0.0229	0.2038	0.6401	0.0532
$10^{1/2}$	100	0.0312	0.2069	0.6190	0.0628
10	100	0.0449	0.1696	0.4499	0.0748

Supplemental Table 2.24: (Experiment I: No Tie) $T = 20$, 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	0.1041	0.1419	0.2704	0.0993
$10^{-1/2}$	2	0.0961	0.1361	0.2579	0.0953
1	2	0.0913	0.1402	0.2656	0.0986
$10^{1/2}$	2	0.0990	0.1243	0.1813	0.1071
10	2	0.1020	0.1072	0.1203	0.1037
10^{-1}	10	0.2446	0.2261	0.5289	0.1309
$10^{-1/2}$	10	0.1975	0.2089	0.5292	0.1103
1	10	0.1336	0.1859	0.4734	0.0948
$10^{1/2}$	10	0.0944	0.1632	0.3754	0.0962
10	10	0.0898	0.1234	0.2125	0.0998
10^{-1}	20	0.3402	0.2655	0.5670	0.1635
$10^{-1/2}$	20	0.2765	0.2383	0.5634	0.1325
1	20	0.1805	0.1879	0.4822	0.0987
$10^{1/2}$	20	0.1078	0.1725	0.4489	0.0884
10	20	0.0901	0.1408	0.2885	0.0965
10^{-1}	50	0.4912	0.3163	0.5988	0.2170
$10^{-1/2}$	50	0.4122	0.2856	0.5864	0.1764
1	50	0.2869	0.2398	0.5807	0.1263
$10^{1/2}$	50	0.1750	0.2012	0.5137	0.1021
10	50	0.1113	0.1811	0.4349	0.1004
10^{-1}	100	0.5995	0.3695	0.7016	0.2539
$10^{-1/2}$	100	0.5088	0.3138	0.6447	0.2055
1	100	0.3653	0.2637	0.6510	0.1456
$10^{1/2}$	100	0.2261	0.2279	0.6444	0.1043
10	100	0.1304	0.1761	0.4686	0.0876

Supplemental Table 2.25: (Experiment II: Exact Tie) $T = 20$, 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	0.1947	0.1005	0.0038	0.1234
10^{-1}	10	0.4617	0.2213	-0.0253	0.2681
$10^{-1/2}$	10	0.4098	0.1712	-0.0737	0.2197
1	10	0.3218	0.1141	-0.0991	0.1534
$10^{1/2}$	10	0.2408	0.0867	-0.0714	0.1131
10	10	0.1940	0.0732	-0.0508	0.0991
10^{-1}	20	0.5589	0.2701	-0.0262	0.3248
$10^{-1/2}$	20	0.5009	0.2214	-0.0653	0.2717
1	20	0.3968	0.1407	-0.1220	0.1900
$10^{1/2}$	20	0.2949	0.1115	-0.0767	0.1378
10	20	0.2272	0.0918	-0.0470	0.1141
10^{-1}	50	0.6826	0.3344	-0.0229	0.4029
$10^{-1/2}$	50	0.6203	0.2693	-0.0907	0.3450
1	50	0.5139	0.2065	-0.1089	0.2649
$10^{1/2}$	50	0.3862	0.1405	-0.1116	0.1831
10	50	0.2824	0.1010	-0.0851	0.1343
10^{-1}	100	0.7606	0.3815	-0.0073	0.4514
$10^{-1/2}$	100	0.6957	0.3089	-0.0880	0.3892
1	100	0.5828	0.2170	-0.1584	0.2969
$10^{1/2}$	100	0.4534	0.1646	-0.1316	0.2160
10	100	0.3322	0.1116	-0.1148	0.1505

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

Supplemental Table 2.26: (Experiment II: Exact Tie) $T = 20$, Variance of the Estimates

μ_*	N	(1) $\hat{\alpha}$	(2) $J(\hat{\alpha})$	(3) $G(\hat{\alpha})$	(4) $\hat{\alpha}_{BC}^{boot}$
**	2	0.0722	0.1352	0.2766	0.0878
10^{-1}	10	0.0340	0.1695	0.4919	0.0591
$10^{-1/2}$	10	0.0363	0.1750	0.4999	0.0622
1	10	0.0450	0.1812	0.4922	0.0724
$10^{1/2}$	10	0.0562	0.1564	0.3786	0.0808
10	10	0.0658	0.1451	0.3228	0.0856
10^{-1}	20	0.0283	0.1925	0.5822	0.0570
$10^{-1/2}$	20	0.0288	0.1801	0.5408	0.0563
1	20	0.0361	0.1867	0.5337	0.0650
$10^{1/2}$	20	0.0479	0.1603	0.4130	0.0745
10	20	0.0586	0.1435	0.3375	0.0795
10^{-1}	50	0.0231	0.2004	0.6242	0.0524
$10^{-1/2}$	50	0.0251	0.2053	0.6351	0.0564
1	50	0.0296	0.2046	0.6163	0.0612
$10^{1/2}$	50	0.0377	0.1763	0.5008	0.0670
10	50	0.0496	0.1640	0.4312	0.0764
10^{-1}	100	0.0188	0.2195	0.7099	0.0473
$10^{-1/2}$	100	0.0196	0.2180	0.7032	0.0483
1	100	0.0227	0.2058	0.6520	0.0521
$10^{1/2}$	100	0.0284	0.1840	0.5518	0.0570
10	100	0.0389	0.1681	0.4652	0.0670

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

Supplemental Table 2.27: (Experiment II: Exact Tie) $T = 20$, MSE of the Estimates

μ_*	N	(1) $\hat{\alpha}$	(2) $J(\hat{\alpha})$	(3) $G(\hat{\alpha})$	(4) $\hat{\alpha}_{BC}^{boot}$
**	2	0.1101	0.1453	0.2766	0.1030
10^{-1}	10	0.2472	0.2185	0.4925	0.1310
$10^{-1/2}$	10	0.2043	0.2043	0.5053	0.1104
1	10	0.1486	0.1942	0.5021	0.0959
$10^{1/2}$	10	0.1142	0.1639	0.2837	0.0936
10	10	0.1025	0.1505	0.3254	0.0954
10^{-1}	20	0.3407	0.2655	0.5829	0.1625
$10^{-1/2}$	20	0.2797	0.2291	0.5451	0.1302
1	20	0.1935	0.2065	0.5486	0.1011
$10^{1/2}$	20	0.1348	0.1727	0.4188	0.0935
10	20	0.1102	0.1520	0.3397	0.0925
10^{-1}	50	0.4890	0.3122	0.6248	0.2148
$10^{-1/2}$	50	0.4099	0.2779	0.6433	0.1754
1	50	0.2937	0.2472	0.6282	0.1314
$10^{1/2}$	50	0.1868	0.1960	0.5132	0.1006
10	50	0.1293	0.1742	0.4384	0.0944
10^{-1}	100	0.5973	0.3651	0.7100	0.2511
$10^{-1/2}$	100	0.5036	0.3134	0.7110	0.1998
1	100	0.3624	0.2529	0.6771	0.1403
$10^{1/2}$	100	0.2340	0.2111	0.5691	0.1036
10	100	0.1493	0.1805	0.4784	0.0897

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

Supplemental Table 2.28: (Experiment III: Near Tie) $T = 20$, 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	0.1906	0.0974	0.0017	0.1193
$10^{-1/2}$	2	0.1820	0.0899	-0.0047	0.1106
1	2	0.1579	0.0705	-0.0193	0.0874
$10^{1/2}$	2	0.1052	0.0337	-0.0397	0.0429
10	2	0.0455	0.0026	-0.0404	0.0073
10^{-1}	10	0.4615	0.2212	-0.0254	0.2679
$10^{-1/2}$	10	0.4091	0.1705	-0.0744	0.2190
1	10	0.3182	0.1106	-0.1023	0.1493
$10^{1/2}$	10	0.2237	0.0686	-0.0905	0.0948
10	10	0.1404	0.0295	-0.0842	0.0483
10^{-1}	20	0.5589	0.2700	-0.0263	0.3248
$10^{-1/2}$	20	0.5006	0.2208	-0.0662	0.2714
1	20	0.3955	0.1389	-0.1242	0.1885
$10^{1/2}$	20	0.2875	0.1022	-0.0879	0.1294
10	20	0.1998	0.0594	-0.0847	0.0867
10^{-1}	50	0.6826	0.3344	-0.0229	0.4029
$10^{-1/2}$	50	0.6203	0.2693	-0.0907	0.3450
1	50	0.5137	0.2057	-0.1102	0.2646
$10^{1/2}$	50	0.3845	0.1384	-0.1141	0.1811
10	50	0.2733	0.0933	-0.0912	0.1241
10^{-1}	100	0.7606	0.3815	-0.0073	0.4514
$10^{-1/2}$	100	0.6957	0.3088	-0.0881	0.3892
1	100	0.5827	0.2167	-0.1588	0.2968
$10^{1/2}$	100	0.4529	0.1640	-0.1325	0.2154
10	100	0.3292	0.1091	-0.1168	0.1471

Supplemental Table 2.29: (Experiment III: Near Tie) $T = 20$, 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	0.0722	0.1341	0.2712	0.0878
$10^{-1/2}$	2	0.0724	0.1349	0.2703	0.0882
1	2	0.0733	0.1336	0.2614	0.0894
$10^{1/2}$	2	0.0785	0.1280	0.2359	0.0936
10	2	0.0914	0.1206	0.1857	0.1021
10^{-1}	10	0.0340	0.1694	0.4915	0.0591
$10^{-1/2}$	10	0.0363	0.1745	0.4987	0.0621
1	10	0.0449	0.1780	0.4837	0.0720
$10^{1/2}$	10	0.0569	0.1572	0.3811	0.0811
10	10	0.0691	0.1500	0.3276	0.0890
10^{-1}	20	0.0283	0.1925	0.5823	0.0570
$10^{-1/2}$	20	0.0288	0.1801	0.5410	0.0563
1	20	0.0360	0.1868	0.5343	0.0649
$10^{1/2}$	20	0.0480	0.1591	0.4098	0.0744
10	20	0.0608	0.1557	0.3683	0.0819
10^{-1}	50	0.0231	0.2004	0.6242	0.0524
$10^{-1/2}$	50	0.0251	0.2054	0.6352	0.0563
1	50	0.0296	0.2051	0.6185	0.0612
$10^{1/2}$	50	0.0377	0.1755	0.4984	0.0669
10	50	0.0499	0.1629	0.4273	0.0764
10^{-1}	100	0.0188	0.2195	0.7100	0.0473
$10^{-1/2}$	100	0.0196	0.2180	0.7033	0.0483
1	100	0.0227	0.2060	0.6525	0.0522
$10^{1/2}$	100	0.0284	0.1843	0.5533	0.0569
10	100	0.0388	0.1659	0.4595	0.0667

Supplemental Table 2.30: (Experiment III: Near Tie) $T = 20$, 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	0.1085	0.1436	0.2712	0.1021
$10^{-1/2}$	2	0.1055	0.1430	0.2704	0.1005
1	2	0.0983	0.1386	0.2618	0.0971
$10^{1/2}$	2	0.0895	0.1292	0.2374	0.0955
10	2	0.0933	0.1206	0.1873	0.1022
10^{-1}	10	0.2470	0.2183	0.4922	0.1309
$10^{-1/2}$	10	0.2037	0.2036	0.5042	0.1100
1	10	0.1461	0.1902	0.4941	0.0943
$10^{1/2}$	10	0.1070	0.1619	0.3893	0.0901
10	10	0.0888	0.1509	0.3347	0.0913
10^{-1}	20	0.3406	0.2655	0.5830	0.1624
$10^{-1/2}$	20	0.2794	0.2289	0.5454	0.1300
1	20	0.1924	0.2061	0.5497	0.1004
$10^{1/2}$	20	0.1307	0.1696	0.4175	0.0912
10	20	0.1007	0.1592	0.3755	0.0894
10^{-1}	50	0.4890	0.3122	0.6247	0.2148
$10^{-1/2}$	50	0.4099	0.2779	0.6434	0.1754
1	50	0.2935	0.2475	0.6307	0.1312
$10^{1/2}$	50	0.1855	0.1947	0.5114	0.0997
10	50	0.1246	0.1716	0.4356	0.0918
10^{-1}	100	0.5973	0.3651	0.7100	0.2511
$10^{-1/2}$	100	0.5036	0.3134	0.7110	0.1998
1	100	0.3623	0.2529	0.6777	0.1402
$10^{1/2}$	100	0.2335	0.2112	0.5709	0.1033
10	100	0.1472	0.1778	0.4731	0.0883

Supplemental Table 2.31: (Experiment I: No Tie) $T = 50$, 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	0.1655	0.0769	-0.0127	0.0913
$10^{-1/2}$	2	0.1368	0.0452	-0.0474	0.0647
1	2	0.0865	0.0228	-0.0415	0.0301
$10^{1/2}$	2	0.0419	0.0036	-0.0352	0.0136
10	2	0.0183	0.0080	-0.0023	0.0075
10^{-1}	10	0.4538	0.2058	-0.0447	0.2549
$10^{-1/2}$	10	0.3937	0.1505	-0.0952	0.1987
1	10	0.2802	0.0742	-0.1339	0.1091
$10^{1/2}$	10	0.1480	0.0076	-0.1342	0.0323
10	10	0.0582	0.0051	-0.0486	0.0030
10^{-1}	20	0.5617	0.2849	0.0053	0.3243
$10^{-1/2}$	20	0.5013	0.2284	-0.0472	0.2707
1	20	0.3821	0.1380	-0.1085	0.1750
$10^{1/2}$	20	0.2327	0.0397	-0.1552	0.0770
10	20	0.1156	0.0175	-0.0815	0.0284
10^{-1}	50	0.6795	0.3383	-0.0064	0.3949
$10^{-1/2}$	50	0.6166	0.2721	-0.0759	0.3379
1	50	0.5022	0.1998	-0.1056	0.2480
$10^{1/2}$	50	0.3571	0.1113	-0.1371	0.1519
10	50	0.2035	0.0566	-0.0917	0.0612
10^{-1}	100	0.7560	0.3936	0.0276	0.4371
$10^{-1/2}$	100	0.6934	0.3274	-0.0423	0.3800
1	100	0.5830	0.2597	-0.0668	0.2950
$10^{1/2}$	100	0.4393	0.1623	-0.1175	0.1966
10	100	0.2864	0.0641	-0.1604	0.1011

Supplemental Table 2.32: (Experiment I: No Tie) $T = 50$, 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	0.0743	0.1576	0.3519	0.0900
$10^{-1/2}$	2	0.0758	0.1646	0.3752	0.0917
1	2	0.0821	0.1424	0.2884	0.0961
$10^{1/2}$	2	0.0906	0.1320	0.2291	0.0992
10	2	0.0948	0.1058	0.1311	0.0981
10^{-1}	10	0.0378	0.2504	0.7629	0.0680
$10^{-1/2}$	10	0.0399	0.2421	0.7280	0.0701
1	10	0.0480	0.2277	0.6688	0.0773
$10^{1/2}$	10	0.0687	0.2036	0.5307	0.0953
10	10	0.0873	0.1394	0.2628	0.1023
10^{-1}	20	0.0301	0.2505	0.7964	0.0612
$10^{-1/2}$	20	0.0309	0.2603	0.8358	0.0613
1	20	0.0363	0.2447	0.7607	0.0663
$10^{1/2}$	20	0.0499	0.2320	0.6830	0.0776
10	20	0.0674	0.1643	0.3958	0.0870
10^{-1}	50	0.0212	0.2923	0.9773	0.0500
$10^{-1/2}$	50	0.0212	0.2979	0.9987	0.0494
1	50	0.0253	0.2609	0.8493	0.0554
$10^{1/2}$	50	0.0349	0.2543	0.7934	0.0652
10	50	0.0545	0.1898	0.5094	0.0818
10^{-1}	100	0.0174	0.2939	1.0040	0.0440
$10^{-1/2}$	100	0.0188	0.3061	1.0381	0.0466
1	100	0.0221	0.2845	0.9516	0.0510
$10^{1/2}$	100	0.0304	0.2590	0.8317	0.0602
10	100	0.0.474	0.2337	0.6835	0.0785

Supplemental Table 2.33: (Experiment I: No Tie) $T = 50$, 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	0.1017	0.1635	0.3521	0.0984
$10^{-1/2}$	2	0.0945	0.1667	0.3775	0.0959
1	2	0.0896	0.1429	0.2901	0.0970
$10^{1/2}$	2	0.0924	0.1320	0.2303	0.0994
10	2	0.0951	0.1058	0.1311	0.0982
10^{-1}	10	0.2438	0.2928	0.7649	0.1330
$10^{-1/2}$	10	0.1949	0.2647	0.7370	0.1096
1	10	0.1265	0.2332	0.6848	0.0892
$10^{1/2}$	10	0.0906	0.2037	0.5487	0.0963
10	10	0.0907	0.1394	0.2652	0.1023
10^{-1}	20	0.3456	0.3317	0.7965	0.1664
$10^{-1/2}$	20	0.2822	0.3125	0.8381	0.1346
1	20	0.1824	0.2637	0.7725	0.0969
$10^{1/2}$	20	0.1040	0.2336	0.7070	0.0836
10	20	0.0807	0.1646	0.4024	0.0878
10^{-1}	50	0.4828	0.4068	0.9773	0.2059
$10^{-1/2}$	50	0.4015	0.3720	1.0045	0.1636
1	50	0.2775	0.3009	0.8605	0.1169
$10^{1/2}$	50	0.1624	0.2666	0.8122	0.0883
10	50	0.0959	0.1930	0.5178	0.0855
10^{-1}	100	0.5890	0.4489	1.0048	0.2351
$10^{-1/2}$	100	0.4995	0.4133	1.0399	0.1910
1	100	0.3620	0.3520	0.9560	0.1380
$10^{1/2}$	100	0.2234	0.2854	0.8455	0.0989
10	100	0.1294	0.2378	0.7093	0.0887

Supplemental Table 2.34: (Experiment II: Exact Tie) $T = 50$, 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	0.1820	0.0943	0.0058	0.1077
10^{-1}	10	0.4571	0.1801	-0.0997	0.2577
$10^{-1/2}$	10	0.4066	0.1573	-0.0945	0.2124
1	10	0.3200	0.1051	-0.1120	0.1487
$10^{1/2}$	10	0.2471	0.1092	-0.0301	0.1218
10	10	0.2066	0.0887	-0.0305	0.1119
10^{-1}	20	0.5628	0.2611	-0.0438	0.3246
$10^{-1/2}$	20	0.5063	0.2244	-0.0605	0.2747
1	20	0.4037	0.1556	-0.0950	0.1969
$10^{1/2}$	20	0.2942	0.1009	-0.0943	0.1349
10	20	0.2240	0.0993	-0.0266	0.1105
10^{-1}	50	0.6803	0.3342	-0.0155	0.3956
$10^{-1/2}$	50	0.6202	0.2906	-0.0423	0.3419
1	50	0.5131	0.2297	-0.0565	0.2606
$10^{1/2}$	50	0.3814	0.1196	-0.1449	0.1730
10	50	0.2745	0.0884	-0.0995	0.1209
10^{-1}	100	0.7569	0.3767	-0.0072	0.4385
$10^{-1/2}$	100	0.6969	0.3278	-0.0450	0.3856
1	100	0.5889	0.2496	-0.0932	0.3017
$10^{1/2}$	100	0.4582	0.1768	-0.1075	0.2166
10	100	0.3321	0.1062	-0.1218	0.1450

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

Supplemental Table 2.35: (Experiment II: Exact Tie) $T = 50$ **, Variance of the Estimates**

μ_*	N	(1) $\hat{\alpha}$	(2) $J(\hat{\alpha})$	(3) $G(\hat{\alpha})$	(4) $\hat{\alpha}_{BC}^{boot}$
**	2	0.0742	0.1572	0.3462	0.0900
10^{-1}	10	0.0384	0.2842	0.8880	0.0695
$10^{-1/2}$	10	0.0409	0.2571	0.7858	0.0726
1	10	0.0494	0.2452	0.7225	0.0824
$10^{1/2}$	10	0.0600	0.1887	0.4997	0.0867
10	10	0.0666	0.1809	0.4589	0.0872
10^{-1}	20	0.0310	0.2822	0.9012	0.0634
$10^{-1/2}$	20	0.0327	0.2628	0.8248	0.0654
1	20	0.0384	0.2568	0.7898	0.0714
$10^{1/2}$	20	0.0496	0.2243	0.6469	0.0799
10	20	0.0579	0.1739	0.4419	0.0807
10^{-1}	50	0.0219	0.3086	1.0319	0.0514
$10^{-1/2}$	50	0.0230	0.2952	0.9821	0.0530
1	50	0.0275	0.2623	0.8435	0.0586
$10^{1/2}$	50	0.0391	0.2774	0.8571	0.0733
10	50	0.0520	0.2252	0.6421	0.0822
10^{-1}	100	0.0173	0.3098	1.0634	0.0438
$10^{-1/2}$	100	0.0181	0.3005	1.0319	0.0448
1	100	0.0224	0.2968	0.9978	0.0517
$10^{1/2}$	100	0.0316	0.2735	0.8756	0.0638
10	100	0.0445	0.2433	0.7281	0.0758

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

Supplemental Table 2.36: (Experiment II: Exact Tie) $T = 50$, MSE of the Estimates

μ_*	N	(1) $\hat{\alpha}$	(2) $J(\hat{\alpha})$	(3) $G(\hat{\alpha})$	(4) $\hat{\alpha}_{BC}^{boot}$
**	2	0.1073	0.1661	0.3463	0.1016
10^{-1}	10	0.2473	0.3167	0.8979	0.1359
$10^{-1/2}$	10	0.2062	0.2819	0.7947	0.1177
1	10	0.1518	0.2563	0.7351	0.1045
$10^{1/2}$	10	0.1210	0.2006	0.5006	0.1015
10	10	0.1093	0.1887	0.4599	0.0997
10^{-1}	20	0.3477	0.3504	0.9031	0.1688
$10^{-1/2}$	20	0.2891	0.3131	0.8285	0.1409
1	20	0.2014	0.2810	0.7989	0.1102
$10^{1/2}$	20	0.1361	0.2345	0.6558	0.0981
10	20	0.1081	0.1838	0.4426	0.0929
10^{-1}	50	0.4847	0.4202	1.0321	0.2079
$10^{-1/2}$	50	0.4076	0.3797	0.9839	0.1699
1	50	0.2907	0.3151	0.8467	0.1265
$10^{1/2}$	50	0.1845	0.2917	0.8781	0.1033
10	50	0.1273	0.2330	0.6520	0.0968
10^{-1}	100	0.5902	0.4517	1.0635	0.2361
$10^{-1/2}$	100	0.5037	0.4079	1.0339	0.1935
1	100	0.3693	0.3591	1.0065	0.1427
$10^{1/2}$	100	0.2415	0.3048	0.8871	0.1107
10	100	0.1548	0.2546	0.7430	0.0969

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

Supplemental Table 2.37: (Experiment III: Near Tie) $T = 50$, 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	0.1796	0.0926	0.0047	0.1054
$10^{-1/2}$	2	0.1745	0.0865	-0.0024	0.1002
1	2	0.1593	0.0698	-0.0205	0.0852
$10^{1/2}$	2	0.1221	0.0391	-0.0448	0.0519
10	2	0.0696	0.0128	-0.0444	0.0214
10^{-1}	10	0.4570	0.1802	-0.0994	0.2576
$10^{-1/2}$	10	0.4061	0.1564	-0.0958	0.2118
1	10	0.3175	0.1021	-0.1155	0.1458
$10^{1/2}$	10	0.2359	0.0995	-0.0384	0.1096
10	10	0.1688	0.0572	-0.0556	0.0740
10^{-1}	20	0.5628	0.2610	-0.0438	0.3246
$10^{-1/2}$	20	0.5062	0.2242	-0.0607	0.2746
1	20	0.4031	0.1563	-0.0930	0.1962
$10^{1/2}$	20	0.2900	0.0967	-0.0986	0.1302
10	20	0.2060	0.0705	-0.0664	0.0915
10^{-1}	50	0.6803	0.3342	-0.0155	0.3956
$10^{-1/2}$	50	0.6202	0.2906	-0.0424	0.3419
1	50	0.5129	0.2295	-0.0567	0.2604
$10^{1/2}$	50	0.3801	0.1180	-0.1467	0.1714
10	50	0.2677	0.0800	-0.1097	0.1129
10^{-1}	100	0.7569	0.3767	-0.0072	0.4385
$10^{-1/2}$	100	0.6969	0.3278	-0.0449	0.3856
1	100	0.5889	0.2493	-0.0936	0.3016
$10^{1/2}$	100	0.4579	0.1759	-0.1089	0.2161
10	100	0.3298	0.1038	-0.1246	0.1424

Supplemental Table 2.38: (Experiment III: Near Tie) $T = 50$, 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	0.0743	0.1550	0.3389	0.0901
$10^{-1/2}$	2	0.0744	0.1565	0.3452	0.0903
1	2	0.0753	0.1625	0.3635	0.0916
$10^{1/2}$	2	0.0798	0.1602	0.3517	0.0966
10	2	0.0913	0.1487	0.2819	0.1057
10^{-1}	10	0.0384	0.2839	0.8870	0.0695
$10^{-1/2}$	10	0.0409	0.2575	0.7871	0.0726
1	10	0.0494	0.2450	0.7234	0.0823
$10^{1/2}$	10	0.0604	0.1842	0.4836	0.0868
10	10	0.0697	0.1726	0.4209	0.0898
10^{-1}	20	0.0310	0.2823	0.9015	0.0634
$10^{-1/2}$	20	0.0327	0.2627	0.8243	0.0654
1	20	0.0384	0.2542	0.7810	0.0713
$10^{1/2}$	20	0.0499	0.2277	0.6587	0.0802
10	20	0.0600	0.1891	0.4901	0.0829
10^{-1}	50	0.0219	0.3086	1.0318	0.0514
$10^{-1/2}$	50	0.0230	0.2952	0.9818	0.0530
1	50	0.0274	0.2619	0.8424	0.0586
$10^{1/2}$	50	0.0391	0.2753	0.8490	0.0733
10	50	0.0523	0.2243	0.6357	0.0825
10^{-1}	100	0.0173	0.3097	1.0634	0.0438
$10^{-1/2}$	100	0.0181	0.3005	1.0320	0.0448
1	100	0.0225	0.2973	0.9996	0.0517
$10^{1/2}$	100	0.0316	0.2734	0.8754	0.0638
10	100	0.0446	0.2429	0.7262	0.0759

Supplemental Table 2.39: (Experiment III: Near Tie) $T = 50$, 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	0.1065	0.1636	0.3389	0.1012
$10^{-1/2}$	2	0.1048	0.1640	0.3452	0.1003
1	2	0.1007	0.1674	0.3639	0.0989
$10^{1/2}$	2	0.0947	0.1618	0.3537	0.0993
10	2	0.0961	0.1489	0.2839	0.1062
10^{-1}	10	0.2472	0.3164	0.8969	0.1358
$10^{-1/2}$	10	0.2058	0.2820	0.7963	0.1175
1	10	0.1502	0.2555	0.7367	0.1035
$10^{1/2}$	10	0.1161	0.1941	0.4851	0.0988
10	10	0.0982	0.1758	0.4240	0.0953
10^{-1}	20	0.3477	0.3504	0.9034	0.1688
$10^{-1/2}$	20	0.2890	0.3129	0.8280	0.1409
1	20	0.2009	0.2787	0.7896	0.1098
$10^{1/2}$	20	0.1340	0.2370	0.6684	0.0971
10	20	0.1024	0.1941	0.4945	0.0913
10^{-1}	50	0.4847	0.4202	1.0321	0.2079
$10^{-1/2}$	50	0.4076	0.3796	0.9836	0.1699
1	50	0.2905	0.3146	0.8456	0.1263
$10^{1/2}$	50	0.1836	0.2892	0.8705	0.1027
10	50	0.1239	0.2307	0.6477	0.0952
10^{-1}	100	0.5901	0.4517	1.0634	0.2361
$10^{-1/2}$	100	0.5037	0.4079	1.0340	0.1935
1	100	0.3692	0.3595	1.0084	0.1427
$10^{1/2}$	100	0.2412	0.3044	0.8873	0.1105
10	100	0.1534	0.2537	0.7417	0.0962

Supplemental Table 2.40: (Experiment I: No Tie) $T = 100$, 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	0.1695	0.0762	-0.0177	0.0942
$10^{-1/2}$	2	0.1404	0.0415	-0.0580	0.0672
1	2	0.0895	0.0303	-0.0293	0.0315
$10^{1/2}$	2	0.0451	0.0175	-0.0102	0.0170
10	2	0.0248	0.0159	0.0070	0.0142
10^{-1}	10	0.4640	0.2064	-0.0526	0.2667
$10^{-1/2}$	10	0.4066	0.1840	-0.0397	0.2146
1	10	0.2970	0.1308	-0.0362	0.1307
$10^{1/2}$	10	0.1625	0.0316	-0.1000	0.0469
10	10	0.0709	0.0198	-0.0315	0.0164
10^{-1}	20	0.5576	0.2625	-0.0342	0.3179
$10^{-1/2}$	20	0.4956	0.2123	-0.0725	0.2615
1	20	0.3764	0.1292	-0.1193	0.1659
$10^{1/2}$	20	0.2289	0.0357	-0.1584	0.0711
10	20	0.0967	-0.0115	-0.1203	0.0013
10^{-1}	50	0.6700	0.3282	-0.0153	0.3783
$10^{-1/2}$	50	0.6109	0.2823	-0.0480	0.3277
1	50	0.4996	0.1941	-0.1129	0.2423
$10^{1/2}$	50	0.3540	0.1097	-0.1359	0.1441
10	50	0.2067	0.0463	-0.1149	0.0588
10^{-1}	100	0.7513	0.3438	-0.0656	0.4244
$10^{-1/2}$	100	0.6898	0.2877	-0.1163	0.3693
1	100	0.5796	0.2135	-0.1545	0.2853
$10^{1/2}$	100	0.4396	0.1704	-0.1002	0.1957
10	100	0.2879	0.0883	-0.1123	0.1046

Supplemental Table 2.41: (Experiment I: No Tie) $T = 100$, 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	0.0705	0.1865	0.4701	0.0859
$10^{-1/2}$	2	0.0731	0.1921	0.4892	0.0890
1	2	0.0816	0.1518	0.3241	0.0959
$10^{1/2}$	2	0.0933	0.1269	0.2174	0.1005
10	2	0.0999	0.1084	0.1323	0.1031
10^{-1}	10	0.0358	0.3330	1.0973	0.0650
$10^{-1/2}$	10	0.0377	0.2942	0.9353	0.0676
1	10	0.0456	0.2399	0.7306	0.0750
$10^{1/2}$	10	0.0643	0.2228	0.6160	0.0895
10	10	0.0833	0.1437	0.2943	0.0961
10^{-1}	20	0.0291	0.3623	1.2233	0.0594
$10^{-1/2}$	20	0.0307	0.3459	1.1639	0.0615
1	20	0.0378	0.3265	1.0631	0.0702
$10^{1/2}$	20	0.0497	0.2791	0.8717	0.0783
10	20	0.0728	0.2062	0.5430	0.0944
10^{-1}	50	0.0207	0.3849	1.3494	0.0483
$10^{-1/2}$	50	0.0223	0.3790	1.3195	0.0505
1	50	0.0284	0.3729	1.2720	0.0604
$10^{1/2}$	50	0.0400	0.3196	1.0399	0.0731
10	50	0.0587	0.2583	0.7655	0.0880
10^{-1}	100	0.0179	0.4360	1.5507	0.0447
$10^{-1/2}$	100	0.0188	0.4158	1.4741	0.0467
1	100	0.0219	0.4145	1.4647	0.0513
$10^{1/2}$	100	0.0282	0.3212	1.0869	0.0585
10	100	0.0422	0.2800	0.9065	0.0720

Supplemental Table 2.42: (Experiment I: No Tie) $T = 100$, 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	0.0993	0.1923	0.4704	0.0948
$10^{-1/2}$	2	0.0928	0.1938	0.4926	0.0936
1	2	0.0896	0.1528	0.3249	0.0969
$10^{1/2}$	2	0.0953	0.1272	0.2175	0.1008
10	2	0.1005	0.1086	0.1323	0.1033
10^{-1}	10	0.2512	0.3756	1.1001	0.1362
$10^{-1/2}$	10	0.2030	0.3281	0.9369	0.1137
1	10	0.1338	0.2570	0.7319	0.0921
$10^{1/2}$	10	0.0908	0.2238	0.6259	0.0917
10	10	0.0883	0.1441	0.2953	0.0964
10^{-1}	20	0.3400	0.4312	1.2245	0.1604
$10^{-1/2}$	20	0.2763	0.3910	1.1691	0.1299
1	20	0.1795	0.3432	1.0773	0.0977
$10^{1/2}$	20	0.1021	0.2804	0.8968	0.0834
10	20	0.0822	0.2064	0.5574	0.0944
10^{-1}	50	0.4697	0.4927	1.3496	0.1914
$10^{-1/2}$	50	0.3955	0.4587	1.3218	0.1579
1	50	0.2780	0.4106	1.2847	0.1191
$10^{1/2}$	50	0.1653	0.3317	1.0583	0.0939
10	50	0.1014	0.2605	0.7787	0.0915
10^{-1}	100	0.5823	0.5542	1.5550	0.2248
$10^{-1/2}$	100	0.4946	0.4986	1.4876	0.1831
1	100	0.3578	0.4601	1.4885	0.1327
$10^{1/2}$	100	0.2214	0.3502	1.0970	0.0968
10	100	0.1251	0.2878	0.9191	0.0829

Supplemental Table 2.43: (Experiment II: Exact Tie) $T = 100$, 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	0.1862	0.0873	-0.0122	0.1113
10^{-1}	10	0.4684	0.2245	-0.0206	0.2717
$10^{-1/2}$	10	0.4197	0.1989	-0.0230	0.2283
1	10	0.3351	0.1362	-0.0637	0.1652
$10^{1/2}$	10	0.2604	0.1160	-0.0293	0.1342
10	10	0.2188	0.1177	0.0160	0.1251
10^{-1}	20	0.5585	0.2624	-0.0352	0.3178
$10^{-1/2}$	20	0.5001	0.2344	-0.0326	0.2650
1	20	0.3979	0.1635	-0.0720	0.1877
$10^{1/2}$	20	0.2927	0.1135	-0.0666	0.1310
10	20	0.2207	0.0949	-0.0315	0.1036
10^{-1}	50	0.6716	0.3436	0.0140	0.3806
$10^{-1/2}$	50	0.6137	0.2607	-0.0941	0.3306
1	50	0.5082	0.2214	-0.0668	0.2504
$10^{1/2}$	50	0.3820	0.1305	-0.1223	0.1698
10	50	0.2799	0.1109	-0.0589	0.1254
10^{-1}	100	0.7512	0.3446	-0.0641	0.4247
$10^{-1/2}$	100	0.6917	0.3228	-0.0480	0.3727
1	100	0.5869	0.2257	-0.1373	0.2942
$10^{1/2}$	100	0.4614	0.1853	-0.0922	0.2157
10	100	0.3423	0.1154	-0.1127	0.1550

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

Supplemental Table 2.44: (Experiment II: Exact Tie) $T = 100$, Variance of the Estimates

μ_*	N	(1) $\hat{\alpha}$	(2) $J(\hat{\alpha})$	(3) $G(\hat{\alpha})$	(4) $\hat{\alpha}_{BC}^{boot}$
**	2	0.0700	0.2014	0.5216	0.0852
10^{-1}	10	0.0357	0.3111	1.0118	0.0643
$10^{-1/2}$	10	0.0371	0.2893	0.9412	0.0653
1	10	0.0447	0.2763	0.8681	0.0741
$10^{1/2}$	10	0.0564	0.2394	0.6940	0.0824
10	10	0.0638	0.1811	0.4765	0.0845
10^{-1}	20	0.0289	0.3698	1.2550	0.0591
$10^{-1/2}$	20	0.0303	0.3233	1.0794	0.0608
1	20	0.0366	0.3046	0.9902	0.0680
$10^{1/2}$	20	0.0468	0.2519	0.7850	0.0741
10	20	0.0606	0.2158	0.5960	0.0833
10^{-1}	50	0.0203	0.3741	1.3142	0.0475
$10^{-1/2}$	50	0.0213	0.4284	1.5061	0.0491
1	50	0.0265	0.3465	1.1800	0.0567
$10^{1/2}$	50	0.0379	0.3123	1.0184	0.0699
10	50	0.0524	0.2648	0.8073	0.0806
10^{-1}	100	0.0170	0.4431	1.5913	0.0426
$10^{-1/2}$	100	0.0169	0.3995	1.4257	0.0417
1	100	0.0203	0.4190	1.4803	0.0480
$10^{1/2}$	100	0.0279	0.3389	1.1418	0.0578
10	100	0.0403	0.3175	1.0317	0.0713

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

Supplemental Table 2.45: (Experiment II: Exact Tie) $T = 100$, MSE of the Estimates

μ_*	N	(1) $\hat{\alpha}$	(2) $J(\hat{\alpha})$	(3) $G(\hat{\alpha})$	(4) $\hat{\alpha}_{BC}^{boot}$
**	2	0.1047	0.2090	0.5217	0.0976
10^{-1}	10	0.2550	0.3615	1.0123	0.1381
$10^{-1/2}$	10	0.2132	0.3288	0.9417	0.1174
1	10	0.1570	0.2949	0.8722	0.1014
$10^{1/2}$	10	0.1242	0.2528	0.6948	0.1004
10	10	0.1117	0.1949	0.4767	0.1001
10^{-1}	20	0.3408	0.4387	1.2562	0.1601
$10^{-1/2}$	20	0.2804	0.3782	1.0805	0.1310
1	20	0.1949	0.3313	0.9954	0.1033
$10^{1/2}$	20	0.1325	0.2648	0.7894	0.0913
10	20	0.1093	0.2248	0.5970	0.0940
10^{-1}	50	0.4714	0.4922	1.3144	0.1923
$10^{-1/2}$	50	0.3980	0.4964	1.5149	0.1584
1	50	0.2847	0.3995	1.1845	0.1194
$10^{1/2}$	50	0.1839	0.3293	1.0334	0.0987
10	50	0.1307	0.2771	0.8108	0.0963
10^{-1}	100	0.5814	0.5618	1.5954	0.2229
$10^{-1/2}$	100	0.4953	0.5037	1.4280	0.1806
1	100	0.3648	0.4700	1.4992	0.1345
$10^{1/2}$	100	0.2409	0.3732	1.1503	0.1043
10	100	0.1575	0.3309	1.0444	0.0953

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

Supplemental Table 2.46: (Experiment III: Near Tie) $T = 100$, 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	0.1845	0.0845	-0.0161	0.1096
$10^{-1/2}$	2	0.1811	0.0832	-0.0151	0.1063
1	2	0.1709	0.0768	-0.0178	0.0965
$10^{1/2}$	2	0.1434	0.0460	-0.0518	0.0714
10	2	0.0932	0.0183	-0.0570	0.0346
10^{-1}	10	0.4683	0.2241	-0.0212	0.2716
$10^{-1/2}$	10	0.4195	0.1979	-0.0248	0.2280
1	10	0.3336	0.1350	-0.0645	0.1636
$10^{1/2}$	10	0.2531	0.1097	-0.0344	0.1262
10	10	0.1932	0.0883	-0.0171	0.0995
10^{-1}	20	0.5584	0.2624	-0.0352	0.3178
$10^{-1/2}$	20	0.5000	0.2346	-0.0321	0.2649
1	20	0.3973	0.1614	-0.0757	0.1871
$10^{1/2}$	20	0.2895	0.1076	-0.0752	0.1272
10	20	0.2072	0.0784	-0.0510	0.0893
10^{-1}	50	0.6716	0.3436	0.0140	0.3806
$10^{-1/2}$	50	0.6137	0.2607	-0.0941	0.3305
1	50	0.5081	0.2211	-0.0673	0.2503
$10^{1/2}$	50	0.3813	0.1275	-0.1276	0.1689
10	50	0.2762	0.1165	-0.0440	0.1214
10^{-1}	100	0.7512	0.3446	-0.0641	0.4247
$10^{-1/2}$	100	0.6917	0.3227	-0.0481	0.3727
1	100	0.5869	0.2257	-0.1373	0.2941
$10^{1/2}$	100	0.4612	0.1845	-0.0935	0.2154
10	100	0.3409	0.1194	-0.1033	0.1533

Supplemental Table 2.47: (Experiment III: Near Tie) $T = 100$, 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	0.0701	0.2060	0.5374	0.0853
$10^{-1/2}$	2	0.0702	0.2000	0.5174	0.0854
1	2	0.0708	0.1934	0.4980	0.0861
$10^{1/2}$	2	0.0742	0.1957	0.5052	0.0899
10	2	0.0854	0.1884	0.4379	0.1011
10^{-1}	10	0.0357	0.3121	1.0158	0.0643
$10^{-1/2}$	10	0.0370	0.2914	0.9496	0.0653
1	10	0.0446	0.2749	0.8651	0.0740
$10^{1/2}$	10	0.0570	0.2377	0.6864	0.0831
10	10	0.0665	0.1946	0.5070	0.0876
10^{-1}	20	0.0289	0.3698	1.2548	0.0591
$10^{-1/2}$	20	0.0303	0.3221	1.0746	0.0608
1	20	0.0366	0.3089	1.0055	0.0680
$10^{1/2}$	20	0.0469	0.2527	0.7887	0.0743
10	20	0.0614	0.2239	0.6266	0.0836
10^{-1}	50	0.0203	0.3741	1.3142	0.0475
$10^{-1/2}$	50	0.0213	0.4284	1.5059	0.0491
1	50	0.0265	0.3467	1.1807	0.0567
$10^{1/2}$	50	0.0380	0.3167	1.0340	0.0700
10	50	0.0528	0.2451	0.7314	0.0809
10^{-1}	100	0.0170	0.4431	1.5913	0.0426
$10^{-1/2}$	100	0.0169	0.3996	1.4262	0.0417
1	100	0.0203	0.4189	1.4801	0.0480
$10^{1/2}$	100	0.0279	0.3404	1.1478	0.0578
10	100	0.0403	0.3046	0.9839	0.0712

Supplemental Table 2.48: (Experiment III: Near Tie) $T = 100$, 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	0.1041	0.2131	0.5376	0.0973
$10^{-1/2}$	2	0.1030	0.2069	0.5176	0.0967
1	2	0.1000	0.1993	0.4984	0.0954
$10^{1/2}$	2	0.0948	0.1979	0.5079	0.0950
10	2	0.0941	0.1887	0.4412	0.1023
10^{-1}	10	0.2549	0.3623	1.0163	0.1380
$10^{-1/2}$	10	0.2130	0.3306	0.9502	0.1172
1	10	0.1559	0.2932	0.8693	0.1008
$10^{1/2}$	10	0.1210	0.2497	0.6876	0.0990
10	10	0.1038	0.2024	0.5073	0.0975
10^{-1}	20	0.3408	0.4386	1.2561	0.1601
$10^{-1/2}$	20	0.2803	0.3771	1.0756	0.1309
1	20	0.1945	0.3349	1.0113	0.1030
$10^{1/2}$	20	0.1307	0.2643	0.7943	0.0905
10	20	0.1043	0.2301	0.6292	0.0916
10^{-1}	50	0.4714	0.4921	1.3144	0.1923
$10^{-1/2}$	50	0.3980	0.4964	1.5148	0.1584
1	50	0.2846	0.3956	1.1852	0.1194
$10^{1/2}$	50	0.1834	0.3330	1.0503	0.0986
10	50	0.1291	0.2587	0.7333	0.0957
10^{-1}	100	0.5814	0.5618	1.5954	0.2229
$10^{-1/2}$	100	0.4953	0.5037	1.4285	0.1806
1	100	0.3647	0.4699	1.4989	0.1345
$10^{1/2}$	100	0.2406	0.3744	1.1565	0.1042
10	100	0.1565	0.3189	0.9946	0.0947

Chapter 3

Estimating Stochastic Frontier Models with Panel Data Using the Split-Sample Jackknife

3.1 Introduction

The aim of this chapter is to consider bias corrections based on the jackknife in stochastic frontier models with panel data. It is well known (e.g., Kim, Kim and Schmidt (2007)) that the usual estimates of technical inefficiency based on fixed effects are biased because the frontier is estimated as the maximum of the firm-specific estimated frontier intercepts, and the "max" operation induces an upward bias in the estimated frontier. If there is no tie for the best firm (in terms of the true frontier intercepts), the bias is of order T^{-1} and the usual panel jackknife of Hahn and Newey (2004) removes the bias of that order. However, Satchachai and Schmidt (2008) showed that, if there is an exact tie for the best firm, the bias is of order $T^{-1/2}$ and a different jackknife (the "generalized panel jackknife") is needed to remove the bias of that order. In either case, if the correct version of the jackknife is used, their simulations indicated that the jackknife did indeed remove bias quite effectively. However, the variance (and therefore the MSE) of the estimators was very large.

Dhaene, Jochmans and Thuybaert (2006) proposed a "split sample jackknife" that was intended to remove bias of order T^{-1} without increasing variance as much as the usual panel jackknife. It is simply two times the original estimate (based on the whole sample) minus the average of the two half-sample estimates. The weights on the various estimates are of much smaller magnitude than for the panel jackknife and as a result its variance is smaller. This estimator is relevant for the case of no tie for the best firm, so that the bias is of order T^{-1}. However, in the case of an exact tie, where the bias is of order $T^{-1/2}$, a different version of the split sample jackknife is needed to remove the leading bias term. We derive that estimator (the "generalized split sample jackknife") in this chapter. We also determine the effects on bias if the wrong version of the jackknife is used; that is, if the split sample jackknife is used but there is an exact tie, or if the generalized split sample jackknife is used but there is not a tie.

Since there appears to be a trade-off between bias and variance, we consider whether there is a split-sample jackknife that is optimal in the sense of minimizing mean square error or variance. For the special case of $N = 2$ firms, we find the optimal weights for the original estimate and the two half-sample estimates. These weights depend on whether there is a tie, and if not, how close we are to having a tie.

We then perform Monte Carlo simulations to evaluate the finite sample relevance of these results, and we compare these to the simulation of Satchachai and Schmidt (2008).

The plan of the chapter is as follows. In Section 3.2, we introduce the model upon which the chapter is used. Section 3.3 describes the split sample jackknife and the generalized split sample jackknife. In Section 3.4, we discuss the "optimal" split sample jackknife for the case of two firms. Section 3.5 reports our simulations. Finally, Section 3.6 contains our conclusion.

3.2 The Model

As in Satchachai and Schmidt (2008), we consider the standard panel data stochastic frontier model with time-invariant technical inefficiency u_i:

$$y_{it} = \alpha_i + x'_{it}\beta + v_{it}, i = 1,\dots,N, t = 1,\dots,T, \tag{3.1}$$

where $\alpha_i = \alpha - u_i$ and $u_i \geq 0$. We consider fixed effect estimation in which the technical inefficiency u_i (and hence α_i) is treated as fixed and no assumption is made about the distribution of u_i or the idiosyncratic errors v_{it}.

Given the within estimate $\hat{\beta}$, the estimates $\hat{\alpha}_i$ are the average 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.

As suggested by Schmidt and Sickles (1984), the estimates of the frontier intercept α and technical inefficiency u_i are as follow

$$\hat{\alpha} = \max_{j=1,\dots,N} \hat{\alpha}_j, \qquad \hat{u}_i = \hat{\alpha} - \hat{\alpha}_i, \qquad i = 1,\dots,N. \tag{3.2}$$

We will regard $\hat{\alpha}$ as an estimate of $\alpha_{(N)} = \max_{j=1,\dots,N} \alpha_j$ and $\hat{u}_i$ as an estimate of $u_i^* = \alpha_{(N)} - \alpha_i = u_i - \min_{j=1,\dots,N} u_j$. Because of the "max" operator in (3.2), $\hat{\alpha}$ is biased upward as an estimate of $\alpha_{(N)}$. That is, the largest $\hat{\alpha}_i$ is more likely to contain positive estimation error than negative error. That is the bias that we wish to remove using the jackknife.

Following Satchachai and Schmidt (2008), we distinguish the following two cases. The first is the case of "no tie," in which there is a unique value of i such that

$\alpha_{(N)} = \alpha_i$ (that is, there is a unique best firm). Hall, Härdle and Simar (1995) show that in this case $\hat{\alpha}$ is asymptotically normal (as $T \to \infty$ with N fixed) and that the bootstrap is valid. In this case the bias of $\hat{\alpha}$ is of order T^{-1}. The second case is the case of an exact tie, so that there are two or more values of i such that $\alpha_{(N)} = \alpha_i$. In this case $\hat{\alpha}$ is not asymptotically normal. The bootstrap is not valid, and Satchachai and Schmidt (2008) show that the bias is of order $T^{-1/2}$.

3.3 Split-Sample Jackknife

In their simulations, Satchachai and Schmidt (2008) found that the panel jackknife and generalized panel jackknife remove most of the bias, but their variances are large. The "split-sample jackknife" proposed by Dhaene, Jochmans and Thuysbaert (2006) is an attempt to remove bias without such a substantial increase in variance.

To describe the split-sample jackknife in a general setting, let the data be indexed by $t = 1,\dots,T$. Let $\hat{\theta}$ be the fixed effects estimator based on all T observations; let $\hat{\theta}^{(1)}$ be the "first-half" sample estimator that omits observations from $t = T/2+1$ through T for all cross-section units and uses only the first $T/2$ observations, and let $\hat{\theta}^{(2)}$ be the "second-half" sample estimator that omits observations from $t = 1$ through $T/2$ for all cross-sectional units and uses only the second $T/2$ observations. Then the split-sample jackknife estimator is

$$SSJ(\hat{\theta}) = 2\hat{\theta} - \frac{1}{2}(\hat{\theta}^{(1)} + \hat{\theta}^{(2)}) = 2\hat{\theta} - 0.5\hat{\theta}^{(1)} - 0.5\hat{\theta}^{(2)}. \tag{3.3}$$

3.3.1 No Tie

This is the case of a unique best firm. The bias of $\hat{\theta}$ (i.e., $\hat{\alpha}$) is of order T^{-1} and we can express it in the following expansion: $E(\hat{\theta}) = \theta + \frac{B}{T} + \frac{D}{T^2} + O(T^{-3})$. The panel jackknife of Hahn and Newey (2004) is

$$J(\hat{\theta}) = T\hat{\theta} + (T-1)\frac{1}{T}\sum_t \hat{\theta}_{(t)}, \tag{3.4}$$

where $\hat{\theta}_{(t)}$ is the delete-observation-t estimator that omits observation t and uses the other $T-1$ observations. It is easy to see that, given the bias expansion above, $E[J(\hat{\theta})] = \theta + O(T^{-2})$ so that the leading term of the bias expansion has been removed.

The split sample jackknife $SSJ(\hat{\theta})$ also removes the leading term of the bias expansion. The motivation is that it might have smaller variance than the panel jackknife, because it uses smaller "weights," for example, with $T = 100$, the panel jackknife multiplies the original estimate by 100 and the mean of the delete-one estimates by 99, while the split sample jackknife multiplies the original estimate by two and the mean of the half-sample estimates by one.

3.3.2 An Exact Tie

When there is a tie for the best firm (the largest α_i), Satchachai and Schmidt (2008) show that the bias of $\hat{\alpha}$ is, for large T, of order $T^{-1/2}$. We express it with the following expansion: $E(\hat{\theta}) = \theta + \frac{B}{\sqrt{T}} + \frac{D}{T} + O(T^{-3/2})$. In this case, Satchachai and Schmidt suggested the generalized panel jackknife:

$$G(\hat{\theta}) = \frac{\sqrt{T}}{\sqrt{T}-\sqrt{T-1}}\hat{\theta} - \frac{\sqrt{T-1}}{\sqrt{T}-\sqrt{T-1}}\frac{1}{T}\sum_t \hat{\theta}_{(t)} . \tag{3.5}$$

This is (even more so then the panel jackknife) an aggressively weighted estimator. (For example, with $T = 100$ the weights in (3.5) are 199.5 and 198.5, versus 100 and 99 for the panel jackknife.) Accordingly its variance is large.

For the case of an exact tie, a split sample jackknife can also be used. We propose the following "generalized split sample jackknife:"

$$GSSJ(\hat{\theta}) = \frac{\sqrt{2}}{\sqrt{2}-1}\hat{\theta} - \frac{1}{\sqrt{2}-1}\frac{1}{2}(\hat{\theta}^{(1)} + \hat{\theta}^{(2)}) \tag{3.6}$$

The weights used here are $\frac{\sqrt{2}}{\sqrt{2}-1} = 3.414$ and $\frac{1}{\sqrt{2}-1} = 2.414$. It is easy to verify that these are the correct weights to remove the leading term $\frac{B}{\sqrt{T}}$ from the bias expansion, and therefore reduce the bias to order T^{-1}. Again, the motivation is that we hope that the variance of this estimator will be smaller than the variance of $G(\hat{\theta})$.

3.3.3 What If The Wrong Version Is Used?

In this section, we show what happens to the bias of the estimate when the wrong version of the split-sample jackknife is used.

First we consider the case that the bias is actually of order $T^{-1/2}$ (there is an exact tie), but we use the split-sample jackknife that is designed to remove bias of order T^{-1}.

Theorem 3.1 If the bias is of order $T^{-1/2}$, the split-sample jackknife corrects the bias by 41.4%.

Proof. We have $E(\hat{\theta}) = \theta + \frac{B}{\sqrt{T}} + \textit{higher order terms}$. So, dropping the higher order terms, we calculate

$$\begin{aligned} E(SSJ(\hat{\theta})) &= 2E(\hat{\theta}) - \frac{1}{2}E(\hat{\theta}^{(1)} + \hat{\theta}^{(2)}) \\ &= 2\left(\theta + \frac{B}{\sqrt{T}}\right) - \frac{1}{2}\left(2\theta + \frac{2B}{\sqrt{T/2}}\right) \\ &= \theta + 0.586\frac{B}{\sqrt{T}}. \end{aligned}$$

Comparing this to the original bias of $\frac{B}{\sqrt{T}}$, we have removed the bias by 41.4%. □

Next we consider the case that the bias is actually of order T^{-1} (there is no tie), but we use the generalized split sample jackknife that is designed to remove bias of order $T^{-1/2}$.

Theorem 3.2 If the bias is of order T^{-1}, the generalized split-sample jackknife reverses the sign and changes the bias by a factor of $\sqrt{2}$.

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

$$\begin{aligned} E(GSSJ(\hat{\theta})) &= \frac{\sqrt{2}}{\sqrt{2}-1}E(\hat{\theta}) - \frac{1}{\sqrt{2}-1}\frac{1}{2}E(\hat{\theta}^{(1)} + \hat{\theta}^{(2)}) \\ &= \frac{\sqrt{2}}{\sqrt{2}-1}\left(\theta + \frac{B}{T}\right) - \frac{1}{\sqrt{2}-1}\frac{1}{2}\left(2\theta + \frac{2B}{T/2}\right) \\ &= \theta - \sqrt{2}\frac{B}{T}. \end{aligned}$$

Now, the bias has a reverse sign and the estimate is overly-corrected by a factor of $\sqrt{2}$. □

3.4 The "Optimal" Split-Sample Jackknife

Both intuition and previous simulations indicate that the jackknife may decrease bias but increase variance. Given such a trade-off, we may wish to consider versions of the jackknife that are optimal in the sense that they minimize MSE (or just variance).

To keep things simple, we will consider estimators that are similar to the split sample jackknife, in the sense that they are linear combinations of the original estimator and the two half-sample estimators. That is, we consider the estimators of the form

$$\tilde{\theta} = a_0\hat{\theta} + a_1\hat{\theta}^{(1)} + a_2\hat{\theta}^{(2)}. \tag{3.7}$$

Then we seek the "optimal" weights a_0, a_1 and a_2. However, instead of focusing only on the weights that correct the bias, we now seek weights that minimize variance or MSE.

We can consider estimators that do, or do not, satisfy the following constraint:

$$a_0 + a_1 + a_2 = 1. \tag{3.8}$$

This is basically the condition for consistency of $\tilde{\theta}$. Note that if we do not impose this constraint, minimization of the variance of $\tilde{\theta}$ is a silly problem since $a_0 = a_1 = a_2 = 0$ is the degenerate solution. But it is a well-posed problem to minimize the variance of $\tilde{\theta}$

subject to this constraint, or to minimize the MSE of $\tilde{\theta}$ either with or without the constraint.

To obtain analytic expressions for a_0, a_1 and a_2, we consider the special case that there are only two firms, i.e., $N = 2$. Although the number of firms considered is restricted, the number of time periods T is not restricted.

We define the following notation

$$a = [a_0 \quad a_1 \quad a_2]' ; \tag{3.9A}$$

$$\hat{\Theta} = [\hat{\theta} \quad \hat{\theta}^{(1)} \quad \hat{\theta}^{(2)}]' ; \tag{3.9B}$$

$$E(\hat{\Theta}) \equiv \Theta = [\Theta_0 \quad \Theta_1 \quad \Theta_2]' ; \tag{3.9C}$$

$$V(\hat{\Theta}) \equiv V = \begin{bmatrix} V_{00} & V_{01} & V_{02} \\ V_{10} & V_{11} & V_{12} \\ V_{20} & V_{21} & V_{22} \end{bmatrix} . \tag{3.9D}$$

We note that $\tilde{\theta} = a'\hat{\Theta}$, so that $E(\tilde{\theta}) = a'\Theta$ and $\text{var}(\tilde{\theta}) = a'Va$.

The optimal estimators that we will derive are infeasible in practice, because they depend on Θ and V. Our interest in them is that we want to see how the optimal weights compare to the weights used by the jackknife procedures of the previous Section. Also, we want to see how much better (in terms of variance or MSE) the optimal estimators are. For example, if the optimal estimator is only slightly better than the original estimator, there will be little point in further exploration of split-sample jackknife procedures.

We now consider the case of a simplified version of the panel data stochastic frontier model. We assume $N = 2$, and we also assume normality and we do not include regressors in the model. So we have

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} \sim N\left[\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}, \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}\right]. \tag{3.10}$$

The object of estimation (θ above) is $\alpha = \max(\alpha_1, \alpha_2)$. Without loss of generality, we take $\alpha_2 = 0$ and $\alpha_1 = \alpha > 0$. Therefore α is what we are trying to estimate, and α/σ is a measure of how close we are to a tie.

The expression for Θ and V, for this simplified model, are derived in Appendix. Given these expressions, we then seek to minimize either of the following quantities

$$\text{var}(\tilde{\alpha}) = a'Va; \tag{3.11A}$$

$$MSE(\tilde{\alpha}) = \text{var}(\tilde{\alpha}) + bias^2(\tilde{\alpha}) = a'Va + a'\Theta\Theta'a - 2\alpha a'\Theta + \alpha^2. \tag{3.11B}$$

3.4.1 Unconstrained "Optimal" Split-Sample Jackknife

The unconstrained minimization (with respect to a) of $\text{var}(\tilde{\alpha})$ is trivial, namely, $a = 0$, $\tilde{\alpha} \equiv 0$ and $\text{var}(\tilde{\alpha}) = 0$. However, the unconstrained minimization of $MSE(\tilde{\alpha})$ is not trivial.

Proposition 3.1 Unconstrained "Optimal" Split-Sample Jackknife. The estimator $\tilde{\alpha} = a'\hat{\Theta}$ that minimizes $MSE(\tilde{\alpha})$ is

$$\tilde{\alpha} = \alpha[(V + \Theta\Theta')^{-1}\Theta]'\hat{\Theta} \tag{3.12}$$

Proof. See Appendix. □

In the case of no tie, the unconstrained "optimal" split-panel jackknife that minimizes MSE is as in (3.12) and is a non-trivial result. In an exact tie case, without loss of generality, we can take $\alpha = 0$. However, this leads us to the trivial solution where $a = 0$ ($a_0 = a_1 = a_2 = 0$) and there is no-trivial solution.

3.4.2 Constrained "Optimal" Split-Sample Jackknife

To maintain the connection with the other versions of jackknife, we impose the consistency constraint $a_0 + a_1 + a_2 = 1$. It is worth mentioning that with the constraint, the estimate may or may not be first-order biased. For example, (i) in the case of no tie, if $a_0 = 2$ and $a_1 = a_2 = -\frac{1}{2}$, i.e., weights of the split sample jackknife, or (ii) in the case of an exact tie, if $a_0 = \frac{\sqrt{2}}{\sqrt{2}-1}$ and $a_1 = a_2 = -\frac{1}{\sqrt{2}-1}$, i.e., weights of the generalized split sample jackknife, then there is no bias of the first-order, in the sense that the leading term in the expansion of the bias has been removed. However, these are the weights that completely remove the first-order bias from the original estimate and are not the choices that minimize variance or MSE.

We can rewrite the constraint as $a_0 = 1 - a_1 - a_2$ and, in matrix notation, as

$a = e_1 + Aa_*$, where $e_1 = [1 \quad 0 \quad 0]'$, $A = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}'$ and $a_* = [a_1 \quad a_2]'$.

Then we can write

$$\text{var}(\tilde{\alpha}) = e_1'Ve_1 + 2a_*'A'Ve_1 + a_*'A'VAa_*, \tag{3.13A}$$

$$MSE(\tilde{\alpha}) = e_1'Ce_1 + 2a_*'A'Ce_1 + a_*'A'CAa_* - 2\alpha e_1'm - 2\alpha a_*'A'm + \alpha^2, \tag{3.13B}$$

where $C = V + \Theta\Theta'$.

Proposition 3.2 Constrained "Optimal" (Minimum MSE) Split-Sample Jackknife. The estimator $\tilde{\alpha}$ that minimizes $MSE(\tilde{\alpha})$ subject to the constraint $a_0 + a_1 + a_2 = 1$ is

$$\tilde{\alpha}(\min MSE) = a_0\hat{\alpha} + a_1\hat{\alpha}^{(1)} + a_2\hat{\alpha}^{(2)},$$

where

$$a_0 = \frac{V_{11} - 2V_{01} - 2(\Theta_1 - \alpha)(\Theta_0 - \Theta_1)}{2V_{00} + V_{11} - 4V_{01} + 2(\Theta_0 - \Theta_1)^2}$$

and

$$a_1 = a_2 = \frac{V_{00} - V_{01} + (\Theta_0 - \alpha)(\Theta_0 - \Theta_1)}{2V_{00} + V_{11} - 4V_{01} + 2(\Theta_0 - \Theta_1)^2} \tag{3.14}$$

Proof. See Appendix. □

If there is no tie for the best firm, the "optimal" weights (3.14) depend on the true variance matrix V and on α (the difference between the two firms' means). For the case of an exact tie, we simply take $\alpha = 0$ and the weights simplify to

$$\tilde{\alpha}(\min MSE) = 1.3767\hat{\alpha} - 0.1887\hat{\alpha}^{(1)} - 0.1880\hat{\alpha}^{(2)}.\text{[2]} \tag{3.15}$$

Corollary 3.1 (i) If the bias is of order $T^{-1/2}$, the "optimal" split-sample jackknife (3.15) corrects the bias by almost 15%; and (ii) if the bias is of order T^{-1}, the "optimal" split-sample jackknife (3.15) corrects the bias by about 38%.

Proof. (i) Suppose $E(\hat{\alpha}) = \alpha + \frac{B}{\sqrt{T}} + \textit{higher order terms}$. Dropping the higher order terms,

[2] The difference between the weights on the half-sample estimates, 0.1887 and 0.1880, is due to the randomness in the Monte Carlo evaluation of V.

$$E(\tilde{\alpha}) = 1.3767\left(\alpha + \frac{B}{\sqrt{T}}\right) - 0.1887\left(\alpha + \frac{B}{\sqrt{T/2}}\right) - 0.1880\left(\alpha + \frac{B}{\sqrt{T/2}}\right)$$
$$= \alpha + 0.8439\frac{B}{\sqrt{T}}.$$

(ii) Suppose $E(\hat{\alpha}) = \alpha + \frac{B}{T} + \textit{higher order terms}$. Dropping the higher order terms,

$$E(\tilde{\alpha}) = 1.3767\left(\alpha + \frac{B}{T}\right) - 0.1887\left(\alpha + \frac{B}{T/2}\right) - 0.1880\left(\alpha + \frac{B}{T/2}\right)$$
$$= \alpha + 0.6233\frac{B}{T}.$$

□

<u>Proposition 3.3</u> Constrained "Optimal" (Minimum Variance) Split-Sample Jackknife. The estimator $\tilde{\alpha}$ that minimizes $\text{var}(\tilde{\alpha})$ subject to the constraint $a_0 + a_1 + a_2 = 1$ is

$$\tilde{\alpha}(\text{min var}) = a_0\hat{\alpha} + a_1\hat{\alpha}^{(1)} + a_2\hat{\alpha}^{(2)},$$

where

$$a_0 = \frac{V_{11} - 2V_{01}}{2V_{00} + V_{11} - 4V_{01}} \text{ and } a_1 = a_2 = \frac{V_{00} - V_{01}}{2V_{00} + V_{11} - 4V_{01}}. \quad (3.16)$$

Proof. See Appendix. □

Note that the "optimal" weights (3.16) only depend on the true variance matrix V. In the case of <u>an exact tie</u>, i.e., $\alpha = 0$, the constrained "optimal" split-sample jackknife simplifies to

$$\tilde{\alpha}(\text{min var}) = 0.5\hat{\alpha} + 0.25\hat{\alpha}^{(1)} + 0.25\hat{\alpha}^{(2)}. \quad (3.17)$$

Corollary 3.2 (i) If the bias is of order $T^{-1/2}$, the bias of the "optimal" split-sample jackknife (3.17) increases by 21%; and (ii) if the bias is of order T^{-1}, the bias of the "optimal" split-sample jackknife (3.17) increases by 50%.

Proof. (i) Suppose $E(\hat{\alpha}) = \alpha + \frac{B}{\sqrt{T}} + \textit{higher order terms}$. Dropping the higher order terms,

$$
\begin{aligned}
E(\tilde{\alpha}(\min \mathrm{var})) &= 0.5\left(\alpha + \frac{B}{\sqrt{T}}\right) + 0.25\left(\alpha + \frac{B}{\sqrt{T/2}}\right) + 0.25\left(\alpha + \frac{B}{\sqrt{T/2}}\right) \\
&= \alpha + 1.2071\frac{B}{\sqrt{T}}.
\end{aligned}
$$

(ii) Suppose $E(\hat{\alpha}) = \alpha + \frac{B}{T} + \textit{higher order terms}$. Dropping the higher order terms,

$$
\begin{aligned}
E(\tilde{\alpha}(\min \mathrm{var})) &= 0.50\left(\alpha + \frac{B}{T}\right) + 0.25\left(\alpha + \frac{B}{T/2}\right) + 0.25\left(\alpha + \frac{B}{T/2}\right) \\
&= \alpha + 1.50\frac{B}{T}.
\end{aligned}
$$

□

As expected, minimizing the variance of the estimator, without regard to bias, will increase bias.

3.5 Simulations

In this section, we investigate the finite sample performance of four estimators: (i) the split-sample jackknife estimate, $SSJ(\hat{\alpha})$; (ii) the generalized split-sample jackknife, $GSSJ(\hat{\alpha})$; (iii) the optimal split-sample jackknife that minimizes the MSE, $\tilde{\alpha}(\min MSE)$;

and (iv) the optimal split-sample jackknife that minimizes the variance, $\tilde{\alpha}(\text{min var})$. In the last two cases, these are the estimators in (3.14) and (3.16) that minimized MSE and variance subject to the restriction $a_0 + a_1 + a_2 = 1$. We also compare the results for these estimators to the original estimator $\hat{\alpha}$, the panel jackknife estimator $J(\hat{\alpha})$, and the generalized panel jackknife $G(\hat{\alpha})$ that were analyzed in Satchachai and Schmidt (2008).

3.5.1 Design of the Experiments

We consider the model with no regressors, as in Satchachai and Schmidt (2008). The inclusion of regressors would not change our results much because the coefficients (β) of the regressors are estimated so much more precisely than the α_i are.

Thus, the data generating process is

$$\begin{aligned} y_{it} &= \alpha + v_{it} - u_i \text{ for } i = 1,\dots,N \text{ and } T = 1,\dots,T \\ &= \alpha_i - v_{it} \end{aligned} \tag{3.18}$$

The u_i are i.i.d. half-normal: $u_i = |U_i|$ where $U_i \sim N(0,\sigma_u^2)$; and the $v_{it} \sim N(0,\sigma_v^2)$. These distributional assumptions are not used in estimation. They just characterize the data generating process.

We employ the parameterization used in Satchachai and Schmidt (2008). Our parameters are N,T and $\mu_* = (\sigma_u)_* / T^{-1/2}\sigma_v$ where $(\sigma_u^2)_* = ((\pi-2)/\pi)\sigma_u^2$. We use μ_* because of its convenient interpretation. It measures the standard deviation of the α_i in units of the standard deviation of the $\hat{\alpha}_i$.

To maintain the connection to Satchachai and Schmidt (2008), we use the same parameter values: we fix $\sigma_v^2/T = 0.1$ and consider $\mu_* = 10^{-1}, 10^{-1/2}, 1, 10^{1/2}$ and 10. Then, for a given value of μ_*, we can determine $(\sigma_u^2)_*$ and σ_u^2, i.e.,

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

For the split-sample jackknife and the generalized split-sample jackknife, we set $T = 10$ and consider sample sizes $N = 2, 10, 20, 50$ and 100.

However, for the "optimal" split-sample jackknife, we set $N = 2$. To calculate the weights a_0, a_1 and a_2, we need values of the parameter $\alpha = \left|\alpha_{(2)} - \alpha_{(1)}\right|$ for a given μ_*. These numbers are shown in Table 3.7.

We consider two experiments with the setup described above:

(1) Experiment I (No Tie). The setup of this experiment is exactly as just described. There are no restrictions on the α_i. They follow from the draws of the half-normal u_i.

(2) Experiment II (An Exact Tie). We generate the data as described above. Now we (the data generator) know which firm is the best firm and the value of its intercept

$\alpha_{(N)} = \max_{j=1,\ldots,N} \alpha_j$. Then, we randomly select one of the other $(N-1)$ firms and set its intercept equal to $\alpha_{(N)}$. Hence, we have created an exact two-way tie for the best firm.

For each configuration of $\{\mu_*, N, T\}$, we perform 1,000 replications. Then we report the bias, variance and MSE for each estimator. The results for $\hat{\alpha}$, $J(\hat{\alpha})$, and $G(\hat{\alpha})$ are taken from Satchachai and Schmidt (2008).

3.5.2 Results

Table 3.1 gives the bias of the various estimators for Experiment I in which there is no tie. When μ_* is small, we are in a sense close to a tie. The generalized jackknife $G(\hat{\alpha})$ and the generalized split-sample jackknife $GSSJ(\hat{\alpha})$ are closed to being unbiased. The panel jackknife $J(\hat{\alpha})$ removes almost half of the bias and the split-sample jackknife removes about 41% of the bias, as theory predicts (Theorem 3.1).

For larger values of μ_*, when we are farther from a tie, the panel jackknife and the split-sample jackknife do a good job of removing bias, while the generalized panel jackknife and the generalized split-sample jackknife overcorrect (reverse the sign of the bias). Again, this is as theory predicts (Theorem 3.2).

Table 3.2 gives the variance of the various estimators. As expected, the split-sample jackknife has a smaller variance than the panel jackknife, and the generalized split-sample jackknife has a smaller variance than the generalized panel jackknife. In all cases, the variance is larger than the original fixed effects estimator $\hat{\alpha}$.

Table 3.3 gives the MSE of the various estimators. The generalized panel jackknife and the generalized split-sample jackknife have large MSEs. The MSE of the split sample jackknife is generally smaller than the MSE of the panel jackknife, but bigger than the MSE of the original fixed effects estimator. However, in some cases (large N and/or small μ_*) the split sample jackknife is better in terms of MSE than the original fixed effects estimator.

Table 3.4, 3.5 and 3.6 give the same results for the case of an exact tie. Generally speaking, the results are similar to those in Table 3.1, 3.2 and 3.3 for $\mu_* = 10^{-1}$ (a near tie). The generalized panel jackknife and the generalized split sample jackknife do a good job of removing bias. The generalized split sample jackknife is better than the generalized panel jackknife, in terms of variance and MSE, but in most cases its MSE is larger than the MSE of the original fixed effects estimate or of the split sample jackknife. The split sample jackknife is better than the original fixed effects estimator, in terms of bias and MSE, for almost all of those exact-tie cases.

Table 3.7 shows the "optimal" weights for the two "optimal" split-sample jackknife estimators, in which MSE and variance are minimized subject to the constraint $a_0 + a_1 + a_2 = 1$. For the estimator that minimizes MSE, for small values of μ_* (and α), the weight a_0 is greater than one, while the weights a_1 and a_2 are negative. These patterns of the "optimal" split-sample jackknife that minimizes MSE are similar to those of other versions of jackknife. On the other hand, all of the weights of the "optimal" split-sample jackknife that minimizes variance are less than one. Clearly smaller weights are helpful in keeping the variance of the estimator small, and this agrees with the fact that the estimators that use large weights to aggressively remove bias, e.g., the generalized panel jackknife, tend to have large variance.

In Table 3.8 we compare the bias, variance and MSE of the "optimal" split-sample jackknife to those of the original fixed effects estimate $\hat{\alpha}$. Comparisons to the other estimators that we have considered can be made by referring to the entries in Tables 3.1-3.6.

First consider $\tilde{\alpha}(\min MSE)$. Its bias is slightly smaller than the bias of the original fixed effects estimator, but its variance is slightly larger. Its MSE is very similar to that of the original estimator. (In fact, in some cases it appears to be slightly larger, which logically cannot be the case. This must be a reflection of numerical inaccuracy, which is small but not small relative to the difference in the MSE of the estimators.) The obvious conclusion is that, while split-sample jackknife methods can be used to remove or reduce bias, they will not be successful in reducing the MSE of the estimate by any meaningful amount.

For $\tilde{\alpha}(\min \text{var})$ the situation is somewhat different. Its variance is slightly smaller than the variance of the original fixed effects estimator, while its bias and MSE are slightly larger. All of these differences are small. So, again, while split-sample jackknife methods can be used to remove or reduce bias, that objective is not compatible with reduction of variance or MSE.

3.6 Conclusions

In this chapter, we have tried to find a jackknife-type estimator of the frontier intercept that has small MSE and/or small variance. We have investigated the performance of the split-sample jackknife estimator and the generalized split-sample jackknife. We also consider the "optimal" split-sample jackknife, which minimizes MSE or variance. In terms of the weights that define the estimators, these estimators are less aggressive in removing bias than the panel jackknife and generalized panel jackknife.

When there is no tie for the best firm, we show that the split-sample jackknife is also effective in removing bias of order T^{-1}, but has smaller variance and smaller MSE than the panel jackknife. Although the "optimal" split-sample jackknife has even smaller variance and MSE, it does not properly remove the bias, i.e., the estimate is still biased upward. Also it is not a feasible estimator outside the simulation setting.

When there is an exact tie, the generalized split-sample jackknife also correctly removes the bias, but again its variance and MSE increase significantly comparing to the original fixed effects estimate. In terms of variance and MSE, it is the worst estimator among the four estimators considered. In this case, the "optimal" split-sample jackknife successfully reduces the variance and MSE. This is not surprising since $\hat{\alpha}$ corresponds to $a_0 = 1, a_1 = a_2 = 0$ and these are not "optimal."

3.7 Output Tables

Table 3.1: (Experiment I: No Tie) $T = 10$, Bias of the Estimates

μ_*	N	$\hat{\alpha}$	$J(\hat{\alpha})$	$SSJ(\hat{\alpha})$	$G(\hat{\alpha})$	$GSSJ(\hat{\alpha})$
10^{-1}	2	0.1671	0.0810	0.0898	-0.0099	-0.0195
$10^{-1/2}$	2	0.1391	0.0522	0.0639	-0.0394	-0.0424
1	2	0.0887	0.0230	0.0274	-0.0463	-0.0594
$10^{1/2}$	2	0.0462	0.0125	0.0130	-0.0230	-0.0340
10	2	0.0294	0.0199	0.0172	0.0100	0.0001
10^{-1}	10	0.4532	0.2113	0.2480	-0.0436	-0.0423
$10^{-1/2}$	10	0.3935	0.1556	0.1921	-0.0951	-0.0927
1	10	0.2809	0.0828	0.1045	-0.1261	-0.1450
$10^{1/2}$	10	0.1504	0.0293	0.0343	-0.0983	-0.1300
10	10	0.0577	-0.0034	-0.0036	-0.0678	-0.0904
10^{-1}	20	0.5566	0.2724	0.3118	-0.0271	-0.0344
$10^{-1/2}$	20	0.4928	0.2093	0.2518	-0.0895	-0.0890
1	20	0.3750	0.1176	0.1593	-0.1537	-0.1457
$10^{1/2}$	20	0.2349	0.0563	0.0762	-0.1321	-0.1483
10	20	0.1136	0.0074	0.0230	-0.1046	-0.1052
10^{-1}	50	0.6699	0.3132	0.3765	-0.0629	-0.0383
$10^{-1/2}$	50	0.6092	0.2678	0.3219	-0.0921	-0.0843
1	50	0.4973	0.1903	0.2352	-0.1334	-0.1354
$10^{1/2}$	50	0.3556	0.1151	0.1426	-0.1385	-0.1586
10	50	0.2059	0.0379	0.0574	-0.1391	-0.1527
10^{-1}	100	0.7584	0.3627	0.4336	-0.0544	-0.0256
$10^{-1/2}$	100	0.6949	0.3127	0.3744	-0.0901	-0.0790
1	100	0.5809	0.2293	0.2805	-0.1413	-0.1443
$10^{1/2}$	100	0.4433	0.1652	0.1956	-0.1280	-0.1548
10	100	0.2950	0.0922	0.1118	-0.1217	-0.1474

Table 3.2: (Experiment I: No Tie) $T = 10$, Variance of the Estimates

μ_*	N	$\hat{\alpha}$	$J(\hat{\alpha})$	$SSJ(\hat{\alpha})$	$G(\hat{\alpha})$	$GSSJ(\hat{\alpha})$
10^{-1}	2	0.0666	0.1130	0.0932	0.2127	0.1768
$10^{-1/2}$	2	0.0696	0.1159	0.0954	0.2149	0.1764
1	2	0.0803	0.1183	0.1020	0.2002	0.1698
$10^{1/2}$	2	0.0932	0.1133	0.1059	0.1582	0.1477
10	2	0.0991	0.1055	0.1042	0.1189	0.1209
10^{-1}	10	0.0355	0.1440	0.0821	0.3871	0.2243
$10^{-1/2}$	10	0.0382	0.1472	0.0861	0.3901	0.2308
1	10	0.0438	0.1478	0.0969	0.3665	0.2406
$10^{1/2}$	10	0.0696	0.1378	0.1098	0.2836	0.2289
10	10	0.0890	0.1282	0.1173	0.2106	0.1997
10^{-1}	20	0.0264	0.1419	0.0705	0.4089	0.2090
$10^{-1/2}$	20	0.0284	0.1467	0.0743	0.4191	0.2182
1	20	0.0359	0.1534	0.0860	0.4191	0.2406
$10^{1/2}$	20	0.0518	0.1388	0.0930	0.3297	0.2236
10	20	0.0757	0.1329	0.1043	0.2613	0.1979
10^{-1}	50	0.0208	0.1500	0.0670	0.4750	0.2107
$10^{-1/2}$	50	0.0215	0.1489	0.0669	0.4518	0.2078
1	50	0.0254	0.1479	0.0697	0.4356	0.2064
$10^{1/2}$	50	0.0359	0.1438	0.0833	0.3896	0.2281
10	50	0.0569	0.1401	0.1002	0.3257	0.2321
10^{-1}	100	0.0191	0.1617	0.0643	0.5014	0.2056
$10^{-1/2}$	100	0.0196	0.1556	0.0655	0.4799	0.2085
1	100	0.0231	0.1587	0.0705	0.4807	0.2156
$10^{1/2}$	100	0.0317	0.1585	0.0802	0.4490	0.2245
10	100	0.0440	0.1400	0.0913	0.3532	0.2316

Table 3.3: (Experiment I: No Tie) $T = 10$, MSE of the Estimates

μ_*	N	$\hat{\alpha}$	$J(\hat{\alpha})$	$SSJ(\hat{\alpha})$	$G(\hat{\alpha})$	$GSSJ(\hat{\alpha})$
10^{-1}	2	0.0945	0.1196	0.1012	0.2128	0.1772
$10^{-1/2}$	2	0.0890	0.1186	0.0995	0.2164	0.1782
1	2	0.0882	0.1188	0.1028	0.2024	0.1733
$10^{1/2}$	2	0.0953	0.1134	0.1061	0.1588	0.1488
10	2	0.1000	0.1059	0.1045	0.1190	0.1209
10^{-1}	10	0.2409	0.1887	0.1436	0.3891	0.2261
$10^{-1/2}$	10	0.1931	0.1715	0.1230	0.3991	0.2394
1	10	0.1272	0.1547	0.1078	0.3824	0.2617
$10^{1/2}$	10	0.0922	0.1386	0.1110	0.2932	0.2458
10	10	0.0923	0.1282	0.1173	0.2152	0.2078
10^{-1}	20	0.3362	0.2162	0.1677	0.4096	0.2101
$10^{-1/2}$	20	0.2712	0.1905	0.1377	0.4271	0.2261
1	20	0.1765	0.1673	0.1113	0.4427	0.2618
$10^{1/2}$	20	0.1070	0.1420	0.0988	0.3472	0.2456
10	20	0.0886	0.1329	0.1048	0.2722	0.2090
10^{-1}	50	0.4696	0.2481	0.2088	0.4610	0.2122
$10^{-1/2}$	50	0.3925	0.2206	0.1705	0.4603	0.2149
1	50	0.2727	0.1841	0.1250	0.4534	0.2247
$10^{1/2}$	50	0.1623	0.1571	0.1037	0.4087	0.2532
10	50	0.0993	0.1415	0.1034	0.3451	0.2554
10^{-1}	100	0.5942	0.2932	0.2524	0.5044	0.2062
$10^{-1/2}$	100	0.5025	0.2534	0.2056	0.4880	0.2148
1	100	0.3605	0.2112	0.1492	0.5006	0.2364
$10^{1/2}$	100	0.2282	0.1858	0.1185	0.4654	0.2485
10	100	0.1310	0.1485	0.1038	0.3680	0.2533

Table 3.4: (Experiment II: Exact Tie) $T = 10$, Bias of the Estimates

μ_*	N	$\hat{\alpha}$	$J(\hat{\alpha})$	$SSJ(\hat{\alpha})$	$G(\hat{\alpha})$	$GSSJ(\hat{\alpha})$
***	2	0.1828	0.0997	0.1054	0.0120	-0.0042
10^{-1}	10	0.4580	0.2165	0.2533	-0.0380	-0.0361
$10^{-1/2}$	10	0.4077	0.1693	0.2065	-0.0820	-0.0780
1	10	0.3261	0.1268	0.1505	-0.0832	-0.0979
$10^{1/2}$	10	0.2499	0.1122	0.1189	-0.0330	-0.0663
10	10	0.2052	0.0847	0.1044	-0.0424	-0.0382
10^{-1}	20	0.5593	0.2696	0.3142	-0.0357	-0.0324
$10^{-1/2}$	20	0.5020	0.2158	0.2608	-0.0859	-0.0805
1	20	0.3988	0.1412	0.1771	-0.1302	-0.1363
$10^{1/2}$	20	0.2938	0.0976	0.1212	-0.1093	-0.1230
10	20	0.2282	0.0954	0.1038	-0.0446	-0.0720
10^{-1}	50	0.6707	0.3104	0.3773	-0.0693	-0.0376
$10^{-1/2}$	50	0.6134	0.2715	0.3279	-0.0889	-0.0758
1	50	0.5124	0.2117	0.2548	-0.1052	-0.1095
$10^{1/2}$	50	0.3926	0.1579	0.1845	-0.0895	-0.1097
10	50	0.2899	0.1237	0.1341	-0.0514	-0.0861
10^{-1}	100	0.7615	0.3702	0.4388	-0.0422	-0.0176
$10^{-1/2}$	100	0.7023	0.3240	0.3874	-0.0747	-0.0580
1	100	0.5950	0.2523	0.3055	-0.1089	-0.1040
$10^{1/2}$	100	0.4665	0.1843	0.2240	-0.1132	-0.1190
10	100	0.3382	0.1236	0.1523	-0.1026	-0.1106

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

Table 3.5: (Experiment II: Exact Tie) $T = 10$, Variance of the Estimates

μ_*	N	$\hat{\alpha}$	$J(\hat{\alpha})$	$SSJ(\hat{\alpha})$	$G(\hat{\alpha})$	$GSSJ(\hat{\alpha})$
***	2	0.0658	0.1107	0.0926	0.2063	0.1767
10^{-1}	10	0.0346	0.1408	0.0798	0.3796	0.2186
$10^{-1/2}$	10	0.0362	0.1411	0.0809	0.3785	0.2180
1	10	0.0424	0.1396	0.0858	0.3598	0.2180
$10^{1/2}$	10	0.0528	0.1210	0.0896	0.2729	0.2028
10	10	0.0620	0.1308	0.0957	0.2785	0.1982
10^{-1}	20	0.0264	0.1449	0.0702	0.4207	0.2069
$10^{-1/2}$	20	0.0282	0.1464	0.0736	0.4187	0.2132
1	20	0.0351	0.1484	0.0850	0.4048	0.2342
$10^{1/2}$	20	0.0471	0.1444	0.0957	0.3596	0.2416
10	20	0.0562	0.1228	0.0958	0.2694	0.2153
10^{-1}	50	0.0208	0.1495	0.0684	0.4570	0.2166
$10^{-1/2}$	50	0.0225	0.1554	0.0723	0.4721	0.2269
1	50	0.0274	0.1554	0.0769	0.4506	0.2292
$10^{1/2}$	50	0.0369	0.1505	0.0870	0.4032	0.2403
10	50	0.0497	0.1385	0.0934	0.3311	0.2245
10^{-1}	100	0.0192	0.1614	0.0643	0.5011	0.2048
$10^{-1/2}$	100	0.0204	0.1594	0.0656	0.4892	0.2048
1	100	0.0247	0.1606	0.0736	0.4784	0.2220
$10^{1/2}$	100	0.0313	0.1551	0.0814	0.4395	0.2339
10	100	0.0430	0.1443	0.0924	0.3680	0.2410

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

Table 3.6: (Experiment II: Exact Tie) $T = 10$, MSE of the Estimates

μ_*	N	$\hat{\alpha}$	$J(\hat{\alpha})$	$SSJ(\hat{\alpha})$	$G(\hat{\alpha})$	$GSSJ(\hat{\alpha})$
***	2	0.0993	0.1207	0.1037	0.2065	0.1767
10^{-1}	10	0.2444	0.1877	0.1440	0.3811	0.2199
$10^{-1/2}$	10	0.2024	0.1698	0.1236	0.3852	0.2241
1	10	0.1487	0.1556	0.1084	0.3667	0.2276
$10^{1/2}$	10	0.1153	0.1336	0.1038	0.2740	0.2072
10	10	0.1041	0.1379	0.1066	0.2803	0.1996
10^{-1}	20	0.3392	0.2176	0.1689	0.4220	0.2079
$10^{-1/2}$	20	0.2802	0.1929	0.1416	0.4260	0.2197
1	20	0.1941	0.1683	0.1163	0.4218	0.2528
$10^{1/2}$	20	0.1334	0.1539	0.1104	0.3716	0.2567
10	20	0.1083	0.1319	0.1066	0.2714	0.2205
10^{-1}	50	0.4706	0.2459	0.2107	0.4618	0.2180
$10^{-1/2}$	50	0.3987	0.2291	0.1798	0.4800	0.2327
1	50	0.2900	0.2002	0.1418	0.4617	0.2412
$10^{1/2}$	50	0.1910	0.1754	0.1211	0.4112	0.2523
10	50	0.1337	0.1538	0.1114	0.3337	0.2319
10^{-1}	100	0.5991	0.2984	0.2569	0.5029	0.2051
$10^{-1/2}$	100	0.5136	0.2644	0.2156	0.4948	0.2082
1	100	0.3787	0.2243	0.1669	0.4903	0.2328
$10^{1/2}$	100	0.2489	0.1891	0.1316	0.4523	0.2481
10	100	0.1573	0.1596	0.1155	0.3785	0.2532

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

Table 3.7: Weight Comparisons between Estimators, $N = 2$ and $T = 10$

Panel 3.7.1: Minimizing MSE

μ_*	α	a_0	a_1	a_2
0	0	1.3767	-0.1887	-0.1880
10^{-1}	0.0034	1.3737	-0.1872	-0.1865
$10^{-1/2}$	0.0107	1.3671	-0.1839	-0.1832
1	0.0337	1.3463	-0.1735	-0.1729
$10^{1/2}$	0.1066	1.2811	-0.1408	-0.1403
10	0.3371	1.0796	-0.0400	-0.0396

Panel 3.7.2: Minimizing Variance

μ_*	α	a_0	a_1	a_2
0	0	0.5000	0.2500	0.2500
10^{-1}	0.0034	0.5039	0.2483	0.2478
$10^{-1/2}$	0.0107	0.5039	0.2483	0.2478
1	0.0337	0.5036	0.2484	0.2480
$10^{1/2}$	0.1066	0.5004	0.2500	0.2496
10	0.3371	0.4656	0.2674	0.2671

Note 1: $\alpha = \frac{1}{NREP} \sum_{repl=1}^{NREP} \left[\alpha_{(2),repl} - \alpha_{(1),repl} \right]$ and $\alpha_{(2),repl} = \max(\alpha_{1,repl}, \alpha_{2,repl})$.

Note 2: Recall,

"Optimal" Split-Sample Jackknife: $a_0\hat{\alpha} + a_1\hat{\alpha}^{(1)} + a_2\hat{\alpha}^{(2)}$

Split-Sample Jackknife: $SSJ(\hat{\alpha}) = 2\hat{\alpha} - 0.5\hat{\alpha}^{(1)} - 0.5\hat{\alpha}^{(2)}$

Generalized Split-Sample Jackknife: $GSSJ(\hat{\alpha}) = 3.414\hat{\alpha} - 1.207\hat{\alpha}^{(1)} - 1.207\hat{\alpha}^{(2)}$

Table 3.8: $N = 2$ **(Restricted), Bias, Variance and MSE of the "Optimal" Split-Sample Jackknife Estimates**

Panel 3.8.1: No Tie

		(1) $\hat{\alpha}$			(2) $\tilde{\alpha}(\min MSE)$			(3) $\tilde{\alpha}(\min \text{var})$		
μ_*	T	Bias	Var	MSE	Bias	Var	MSE	Bias	Var	MSE
10^{-1}	10	0.1671	0.0666	0.0945	0.1382	0.0734	0.0925	0.2055	0.0634	0.1057
$10^{-1/2}$	10	0.1391	0.0696	0.0890	0.1115	0.0761	0.0885	0.1764	0.0665	0.0976
1	10	0.0887	0.0803	0.0882	0.0674	0.0854	0.0899	0.1192	0.0776	0.0918
$10^{1/2}$	10	0.0462	0.0932	0.0953	0.0369	0.0953	0.0967	0.0628	0.0921	0.0960
10	10	0.0294	0.0991	0.1000	0.0284	0.0993	0.1001	0.0358	0.0987	0.1000

Panel 3.8.2: Exact Tie

		(1) $\hat{\alpha}$			(2) $\tilde{\alpha}(\min MSE)$			(3) $\tilde{\alpha}(\min \text{var})$		
μ_*	T	Bias	Var	MSE	Bias	Var	MSE	Bias	Var	MSE
***	10	0.1828	0.0658	0.0993	0.1536	0.0727	0.0963	0.2216	0.0626	0.1117

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

3.8 Appendix: Deriving The Expected Value and Variance of The Max

Consider a bivariate normal ($N = 2$) with $t = 1,\dots,T$

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} \sim N\left[\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}, \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}\right]. \tag{3.19}$$

Suppose that $\alpha_1 = \alpha > 0$ and $\alpha_2 = 0$. Without loss of generality, we can assume $T = 2$. All that we are interested in is, for each variable, the overall sample mean and the two half-sample means. So, for each variable, we have effectively two observations (the two half-sample means) and their mean value (since the overall sample mean is indeed the average of the half-sample means).

We are interested in three estimators:

(1) $\hat{\alpha} = \max(\bar{y}_1, \bar{y}_2)$, where $\bar{y}_1 = \dfrac{y_{11} + y_{12}}{2}$ and $\bar{y}_2 = \dfrac{y_{21} + y_{22}}{2}$;

(2) $\hat{\alpha}^{(1)} = \max(y_{11}, y_{21})$; and

(3) $\hat{\alpha}^{(2)} = \max(y_{12}, y_{22})$.

With an approach similar to Satchachai and Schmidt (2008), we can derive the first and second moment of these three estimators:

$$\begin{aligned} E(\hat{\alpha}) &= E[\max(\bar{y}_1, \bar{y}_2)] \\ &= P(\bar{y}_1 \geq \bar{y}_2) \cdot E(\bar{y}_1 \mid \bar{y}_1 \geq \bar{y}_2) + P(\bar{y}_2 \geq \bar{y}_1) \cdot E(\bar{y}_2 \mid \bar{y}_2 \geq \bar{y}_1) \end{aligned} \tag{3.20}$$

and, for $t = 1,2$,

$$\begin{aligned} E(\hat{\alpha}^{(1)}) &= E(\hat{\alpha}^{(2)}) \\ &= E[\max(y_{1t}, y_{2t})] \\ &= P(y_{1t} \geq y_{2t}) \cdot E(y_{1t} \mid y_{1t} \geq y_{2t}) + P(y_{2t} \geq y_{1t}) \cdot E(y_{2t} \mid y_{2t} \geq y_{1t}). \end{aligned} \tag{3.21}$$

From basic probability theory,

$$\begin{aligned}
P(\bar{y}_1 \geq \bar{y}_2) &= P(\bar{y}_1 - \bar{y}_2 \geq 0) \\
&= P\left(\frac{(\bar{y}_1 - \alpha) - (\bar{y}_2 - 0)}{\sigma} \geq -\frac{(\alpha - 0)}{\sigma}\right) \\
&= 1 - \Phi\left(-\frac{\alpha}{\sigma}\right) \\
&= \Phi\left(\frac{\alpha}{\sigma}\right).
\end{aligned}$$

Similarly,

$$P(\bar{y}_2 \geq \bar{y}_1) = \Phi\left(-\frac{\alpha}{\sigma}\right),\ P(y_{1t} \geq y_{2t}) = \Phi\left(\frac{\alpha}{\sqrt{2\sigma^2}}\right), \text{and } P(y_{2t} \geq y_{1t}) = \Phi\left(-\frac{\alpha}{\sqrt{2\sigma^2}}\right).$$

We use the fact about the incidentally truncated bivariate normal:

$$\begin{aligned}
E(\bar{y}_1 \mid \bar{y}_1 \geq \bar{y}_2) &= E(\bar{y}_1 \mid \bar{y}_1 - \bar{y}_2 \geq 0) \\
&= E\left[\bar{y}_1 \mid \frac{(\bar{y}_1 - \alpha) - (\bar{y}_2 - 0)}{\sigma} \geq -\frac{(\alpha - 0)}{\sigma}\right] \\
&= \alpha + \frac{1}{\sqrt{2}}\frac{\sigma}{\sqrt{2}}\lambda\left(-\frac{\alpha}{\sigma}\right); \text{since } \rho = \frac{1}{\sqrt{2}} \\
&= \alpha + \frac{\sigma}{2}\lambda\left(-\frac{\alpha}{\sigma}\right),
\end{aligned}$$

where $\lambda(x) = \frac{\phi(x)}{1 - \Phi(x)}$ or the "inverse Mill's ratio." Similarly,

$E(\bar{y}_2 \mid \bar{y}_2 \geq \bar{y}_1) = \frac{\sigma}{2}\lambda\left(\frac{\alpha}{\sigma}\right)$. Then,

$$\begin{aligned}
E[\max(\bar{y}_1, \bar{y}_2)] &= \Phi\left(\frac{\alpha}{\sigma}\right)\left[\alpha + \frac{\sigma}{2}\lambda\left(-\frac{\alpha}{\sigma}\right)\right] + \Phi\left(-\frac{\alpha}{\sigma}\right)\left[\frac{\sigma}{2}\lambda\left(\frac{\alpha}{\sigma}\right)\right] \\
&= \alpha \cdot \Phi\left(\frac{\alpha}{\sigma}\right) + \sigma \cdot \phi\left(\frac{\alpha}{\sigma}\right)
\end{aligned}$$

and, for $t = 1, 2$,

$$E[\max(y_{1t},y_{2t})] = \Phi\left(\frac{\alpha}{\sqrt{2\sigma^2}}\right)\left[\alpha + \frac{\sigma}{\sqrt{2}}\lambda\left(-\frac{\alpha}{\sqrt{2\sigma^2}}\right)\right] + \Phi\left(-\frac{\alpha}{\sqrt{2\sigma^2}}\right)\left[\frac{\sigma}{\sqrt{2}}\lambda\left(\frac{\alpha}{\sqrt{2\sigma^2}}\right)\right]$$

$$= \alpha\cdot\Phi\left(\frac{\alpha}{\sqrt{2\sigma^2}}\right) + \sqrt{2}\sigma\cdot\phi\left(\frac{\alpha}{\sqrt{2\sigma^2}}\right).$$

Now, we show the derivation of the variance of $\hat{\alpha}, \hat{\alpha}^{(1)}$, and $\hat{\alpha}^{(2)}$. From the definition of variance, we can write

$$\mathrm{var}(\hat{\alpha}) \equiv \mathrm{var}[\max(\bar{y}_1,\bar{y}_2)] = E[\max(\bar{y}_1,\bar{y}_2)]^2 - [E[\max(\bar{y}_1,\bar{y}_2)]]^2 \quad (3.22A)$$

and

$$\begin{aligned}\mathrm{var}(\hat{\alpha}^{(1)}) = \mathrm{var}(\hat{\alpha}^{(2)}) &\equiv \mathrm{var}[\max(y_{1t},y_{2t})] \\ &= E[\max(y_{1t},y_{2t})]^2 - [E[\max(y_{1t},y_{2t})]]^2.\end{aligned} \quad (3.22B)$$

With a similar approach to the one used above, we can also derive $E[\max(\bar{y}_1,\bar{y}_2)]^2$ and $E[\max(y_{1t},y_{2t})]^2$:

$$E[\max(\bar{y}_1,\bar{y}_2)]^2 = P(\bar{y}_1 \ge \bar{y}_2)\cdot E(\bar{y}_1^2 \mid \bar{y}_1 \ge \bar{y}_2) + P(\bar{y}_2 \ge \bar{y}_1)\cdot E(\bar{y}_2^2 \mid \bar{y}_2 \ge \bar{y}_1) \quad (3.23A)$$

and

$$\begin{aligned}E[\max(y_{1t},y_{2t})]^2 = &P(y_{1t} \ge y_{2t})\cdot E(y_{1t}^2 \mid y_{1t} \ge y_{2t}) + \\ &P(y_{2t} \ge y_{1t})\cdot E(y_{2t}^2 \mid y_{2t} \ge y_{1t}).\end{aligned} \quad (3.23B)$$

From the fact about the variance of incidental truncated bivariate normal,

$$\begin{aligned}\mathrm{var}(x \mid truncation) &= \sigma_x^2(1-\rho^2\delta(\alpha_y)) \\ &= E(x^2 \mid truncation) - [E(x \mid truncation)]^2,\end{aligned} \quad (3.24)$$

where $\delta(x) = \lambda(x)[\lambda(x) - x]$. So,

$$E(x^2 \mid truncation) = \mathrm{var}(x \mid truncation) + [E(x \mid truncation)]^2$$

and

$$
\begin{aligned}
E(\bar{y}_1^2 \mid \bar{y}_1 \geq \bar{y}_2) &= \operatorname{var}(\bar{y}_1 \mid \bar{y}_1 \geq \bar{y}_2) + [E(\bar{y}_1 \mid \bar{y}_1 \geq \bar{y}_2)]^2 \\
&= \frac{\sigma^2}{2}\left[1 - \frac{1}{2}\delta\left(-\frac{\alpha}{\sigma}\right)\right] + \left[\alpha + \frac{\sigma}{2}\lambda\left(-\frac{\alpha}{\sigma}\right)\right]^2 \\
&= \frac{\sigma^2}{2} - \frac{\sigma^2}{4}\delta\left(-\frac{\alpha}{\sigma}\right) + \alpha^2 + \frac{\sigma^2}{4}\lambda^2\left(-\frac{\alpha}{\sigma}\right) + \alpha\sigma\lambda\left(-\frac{\alpha}{\sigma}\right)
\end{aligned}
$$

and

$$
\begin{aligned}
E(\bar{y}_2{}^2 \mid \bar{y}_2 \geq \bar{y}_1) &= \operatorname{var}(\bar{y}_2 \mid \bar{y}_2 \geq \bar{y}_1) + [E(\bar{y}_2 \mid \bar{y}_2 \geq \bar{y}_1)]^2 \\
&= \frac{\sigma^2}{2}\left[1 - \frac{1}{2}\delta\left(\frac{\alpha}{\sigma}\right)\right] + \left[\frac{\sigma}{2}\lambda\left(\frac{\alpha}{\sigma}\right)\right]^2 \\
&= \frac{\sigma^2}{2} - \frac{\sigma^2}{4}\delta\left(\frac{\alpha}{\sigma}\right) + \frac{\sigma^2}{4}\lambda^2\left(\frac{\alpha}{\sigma}\right).
\end{aligned}
$$

Then,

$$
\begin{aligned}
E[\max(\bar{y}_1, \bar{y}_2)]^2 &= P(\bar{y}_1 \geq \bar{y}_2) \cdot E(\bar{y}_1^2 \mid \bar{y}_1 \geq \bar{y}_2) + P(\bar{y}_2 \geq \bar{y}_1) \cdot E(\bar{y}_2^2 \mid \bar{y}_2 \geq \bar{y}_1) \\
&= \Phi\left(\frac{\alpha}{\sigma}\right)\left\{\frac{\sigma^2}{2} - \frac{\sigma^2}{4}\delta\left(-\frac{\alpha}{\sigma}\right) + \alpha^2 + \frac{\sigma^2}{4}\lambda^2\left(-\frac{\alpha}{\sigma}\right) + \alpha\sigma\lambda\left(-\frac{\alpha}{\sigma}\right)\right\} \\
&\quad + \Phi\left(-\frac{\alpha}{\sigma}\right)\left\{\frac{\sigma^2}{2} - \frac{\sigma^2}{4}\delta\left(-\frac{\alpha}{\sigma}\right) + \frac{\sigma^2}{4}\lambda^2\left(-\frac{\alpha}{\sigma}\right)\right\}.
\end{aligned}
$$

Substitute $\lambda(x) = \dfrac{\phi(x)}{1 - \Phi(x)}$ and $\delta(x) = \lambda(x)[\lambda(x) - x]$, and we get

$$
\begin{aligned}
E[\max(\bar{y}_1, \bar{y}_2)]^2 &= \frac{\sigma^2}{2} + \alpha^2\Phi\left(\frac{\alpha}{\sigma}\right) + \alpha\sigma\phi\left(\frac{\alpha}{\sigma}\right) \\
&= \frac{\sigma^2}{2} + \alpha \cdot \left[\alpha\Phi\left(\frac{\alpha}{\sigma}\right) + \sigma\phi\left(\frac{\alpha}{\sigma}\right)\right] \\
&= \frac{\sigma^2}{2} + \alpha \cdot E[\max(\bar{y}_1, \bar{y}_2)].
\end{aligned}
$$

and

$$\begin{aligned}\text{var}[\max(\bar{y}_1,\bar{y}_2)] &= E[\max(\bar{y}_1,\bar{y}_2)]^2 - [E[\max(\bar{y}_1,\bar{y}_2)]]^2 \\ &= \frac{\sigma^2}{2} + \alpha \cdot E[\max(\bar{y}_1,\bar{y}_2)] - [E[\max(\bar{y}_1,\bar{y}_2)]]^2 \\ &= \frac{\sigma^2}{2} + E[\max(\bar{y}_1,\bar{y}_2)] \cdot \{\alpha - E[\max(\bar{y}_1,\bar{y}_2)]\}.\end{aligned}$$

So, $\text{var}(\hat{\alpha}) = \frac{\sigma^2}{2} + E(\hat{\alpha}) \cdot (-bias(\hat{\alpha}))$. In a similar manner, the variances of $\hat{\alpha}^{(1)}$ and $\hat{\alpha}^{(2)}$ are $\text{var}(\hat{\alpha}^{(1)}) = \sigma^2 + E(\hat{\alpha}^{(1)}) \cdot (-bias(\hat{\alpha}^{(1)}))$ and $\text{var}(\hat{\alpha}^{(2)}) =$ $= \sigma^2 + E(\hat{\alpha}^{(2)}) \cdot (-bias(\hat{\alpha}^{(2)}))$, respectively.

3.9 Appendix: The "Optimal" Split-Sample Jackknife

Consider a new estimator $\tilde{\alpha}$ that is a linear combination of the estimators based on the full sample, $\hat{\alpha}$, and the half samples, $\hat{\alpha}^{(1)}$ and $\hat{\alpha}^{(2)}$:

$$\tilde{\alpha} = a_0\hat{\alpha} + a_1\hat{\alpha}^{(1)} + a_2\hat{\alpha}^{(2)}, \qquad (3.25)$$

where

(1) $\hat{\alpha} = \max(\bar{y}_1,\bar{y}_2)$, where $\bar{y}_1 = \frac{y_{11}+y_{12}}{2}$ and $\bar{y}_2 = \frac{y_{21}+y_{22}}{2}$;

(2) $\hat{\alpha}^{(1)} = \max(y_{11}, y_{21})$; and

(3) $\hat{\alpha}^{(2)} = \max(y_{12}, y_{22})$.

We are interested in the estimator $\tilde{\alpha}$ that minimizes (i) $MSE(\tilde{\alpha})$, (ii) $MSE(\tilde{\alpha})$ subject to $a_0 + a_1 + a_2 = 1$, and (iii) $\text{var}(\tilde{\alpha})$ subject to $a_0 + a_1 + a_2 = 1$.

First, we define the following notation

$$a = [a_0 \quad a_1 \quad a_2]'; \qquad (3.26A)$$

$$\hat{\Theta} = [\hat{\alpha} \quad \hat{\alpha}^{(1)} \quad \hat{\alpha}^{(2)}]' ; \tag{3.26B}$$

$$E(\hat{\Theta}) \equiv \Theta = [\Theta_0 \quad \Theta_1 \quad \Theta_2]; \text{ and} \tag{3.26C}$$

$$V(\hat{\Theta}) \equiv V = \begin{bmatrix} V_{00} & V_{01} & V_{02} \\ V_{10} & V_{11} & V_{12} \\ V_{20} & V_{21} & V_{22} \end{bmatrix}. \tag{3.26D}$$

Since we assume independence, $V_{12} = V_{21} = 0$. Using matrix algebra,

$bias(\tilde{\alpha}) = a'\Theta - \alpha$, $\text{var}(\tilde{\alpha}) = a'Va$ and $MSE(\tilde{\alpha}) = \text{var}(\tilde{\alpha}) + bias^2(\tilde{\alpha}) =$

$= a'Ca - 2\alpha a'\Theta + \alpha^2$, where $C = V + \Theta\Theta'$. For the constrained cases, we can rewrite $a_0 + a_1 + a_2 = 1$ as $a_0 = 1 - a_1 - a_2$. In matrix notation, we can write $a = e_1 + Aa_*$,

where $e_1 = [1 \quad 0 \quad 0]'$, $A = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}'$ and $a_* = [a_1 \quad a_2]'$.

With the matrix simplification, we can rewrite

$$\begin{aligned} \text{var}(\tilde{\alpha}) &= (e_1 + Aa_*)'V(e_1 + Aa_*) \\ &= e_1'Ve_1 + 2a_*'A'Ve_1 + a_*'A'VAa_* \end{aligned}$$

and

$$\begin{aligned} MSE(\tilde{\alpha}) &= (e_1 + Aa_*)'C(e_1 + Aa_*) - 2\alpha(e_1 + Aa_*)'\Theta + \alpha^2 \\ &= e_1'Ce_1 + 2a_*'A'Ce_1 + a_*'A'CAa_* - 2\alpha e_1'\Theta - 2\alpha a_*'A'\Theta + \alpha^2. \end{aligned}$$

3.9.1: $\tilde{\alpha}$ that minimizes $MSE(\tilde{\alpha})$ without constraint.

F.O.C

$$\frac{dMSE(\tilde{\alpha})}{da} = 2(V + \Theta\Theta')a - 2\alpha\Theta = 0$$

$$(V + \Theta\Theta')a = \alpha\Theta$$

$$a = \alpha \cdot (V + \Theta\Theta')^{-1}\Theta.$$

3.9.2: $\tilde{\alpha}$ that minimizes $MSE(\tilde{\alpha})$ subject to $a_0 + a_1 + a_2 = 1$.

F.O.C

$$\frac{dMSE(\tilde{\alpha})}{da_*} = 2A'Ce_1 + 2A'CAa_* - 2\alpha A'\Theta = 0$$

$$A'CAa_* = \alpha A'\Theta - A'Ce_1$$

$$a_* = (A'CA)^{-1}(\alpha A'\Theta - A'Ce_1).$$

So,

$$a_1 = a_2 = \frac{V_{00} - V_{01} + (\alpha - \Theta_0)(\Theta_1 - \Theta_0)}{2V_{00} + V_{11} - 4V_{01} + 2(\Theta_1 - \Theta_0)^2}$$

and

$$a_0 = 1 - a_1 - a_2 = \frac{V_{11} - 2V_{01} - 2(\Theta_1 - \alpha)(\Theta_0 - \Theta_1)}{2V_{00} + V_{11} - 4V_{01} + 2(\Theta_0 - \Theta_1)^2}.$$

3.9.3: $\tilde{\alpha}$ that minimizes $\text{var}(\tilde{\alpha})$ subject to $a_0 + a_1 + a_2 = 1$.

F.O.C

$$\frac{d\,\text{var}(\tilde{\alpha})}{da_*} = 2A'Ve_1 + 2A'VAa_* = 0$$

$$A'VAa_* = -A'Ve_1$$

$$a_* = -(A'VA)^{-1}A'Ve.$$

So $a_1 = a_2 = \dfrac{V_{00} - V_{01}}{2V_{00} + V_{11} - 4V_{01}}$ and $a_0 = 1 - a_1 - a_2 = \dfrac{V_{11} - 2V_{01}}{2V_{00} + V_{11} - 4V_{01}}$.

BIBLIOGRAPHY

Aigner, D., C.A. K. Lovell and P. Schmidt, 1977, Formulation and Estimation of Stochastic Frontier Production Function Models, *Journal of Econometrics*, 6, 21-37.

Clark, C.E., 1961, The Greatest of a Finite Set of Random Variables, *Operations Research*, 9, 145-161.

Coelli, T, 1995, Estimators and Hypothesis Tests for a Stochastic Frontier Function: A Monte Carlo Analysis, *Journal of Productivity Analysis*, 6, 247-268.

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.

Doran, H. and P. Schmidt, 2006, GMM Estimators with Improved Finite Sample Properties Using Principal Components of the Weighting Matrix with Application to the Dynamic Panel Data Model, *Journal of Econometrics*, 133, 387-409.

Efron, B., 1982, The Jackknife, the Bootstrap and Other Resampling Plans, Society for Industrial Applied mathematics, Philadelphia, Pennsylvania.

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.

Gray, H.L., W.R. Schucany, and T.A. Watkins, 1975, On the Generalized Jackknife and its Relation to Statistical Differentials, *Biometrika*, 63, 637-642.

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

Han, C., L. Orea and P. Schmidt, 2005, Estimation of a Panel Data Model with Parametric Temporal Variation in Individual Effects, *Journal of Econometrics*, 126, 241-267.

Johnson, N.L. and S. Kotz, 1972, Distributions in Statistics: Continuous Multivariate Distributions, Johnson Wiley & Sons, Inc., New York.

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.

Mátyás, L., 1999, Generalized Methods of Moments Estimation, Cambridge University Press, Cambridge; New York.

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

Olson, J.A., P. Schmidt and D.M. Waldman, 1980, A Monte Carlo Study of Estimators of Stochastic Frontier Production Functions, *Journal of Econometrics*, 13, 67-82.

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.

Satchachai, P. and P. Schmidt, 2007, GMM with More Moment Conditions than Observations, *Economics Letters*, 99, 272-275.

-----------------------------------, 2008, Estimates of Technical Inefficiency in Stochastic Frontier Models with Panel Data: Generalized Panel Jackknife Estimation, Department of Economics, Michigan State University.

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 Sample, (Abstract), *Annals of Mathematical Statistics*, 18, 614.