Shrinkage Estimation of the Dose for a Given Mortality

by

S. E. Ahmed and C. J. Nicol
Department of Mathematics and Statistics
and Department of Economics
University of Regina
Regina, Saskatchewan, S4S 0A2
CANADA

Internet: Chris.Nicol@URegina.CA;
http://www.econ.uregina.ca/nicolc

August 19, 2000

Ahmed acknowledges support from the Natural Sciences and Engineering Research Council of Canada. Nicol acknowledges support from the Social Sciences and Humanities Research Council of Canada.

Abstract

This article considers the estimation of the dose for a given mortality rate in a logit model when additional information on that dose is available in the form of a realistic conjecture, perhaps based on previous experience with the drug being tested, or a drug with similar properties. Three classes of point estimation, namely, the unrestricted estimator, the shrinkage estimator and the shrinkage preliminary test estimator are proposed. The asymptotic bias and mean squared errors are derived for each estimator, and are compared. The relative dominance relationship of the estimators is also presented. Interestingly, the proposed shrinkage preliminary test estimator dominates the unrestricted estimator in a range that is wider than that of the usual preliminary test estimator. The results of a simulation experiment indicate that the shrinkage preliminary test estimator yields superior performance over an important range of the parameter space, insofar as applied implementation of the procedure is concerned.

Key Words and Phrases: ED_{100p}, *shrinkage estimator; shrinkage preliminary test estimator; asymptotic biases and risks; asymptotic efficiency.*

1 Introduction

In many situations there is a need to estimate ED_{100p} values, which are doses which correspond, under a given model, to $100p\%$ mortality. For example ED_{50} is the dose corresponding to 50% mortality, a commonly used summary of a fitted model in quantal assay. For notational convenience, ED_{100p} will be denoted by μ_{p_i}. Here, i indexes the proportion associated with the i'th dose level.

In the present investigation, we propose large-sample estimation methods for the dose for a given mortality (μ_p) in a logit model, incorporating the uncertain prior information (*UPI*) that $\mu_{p_i} = \mu_{p_0}$, which may be available to the experimenter. Finite sample studies on this subject have been carried out by Hamilton (1979) and Hoekstra (1991), amongst others.

Our main interest here is in estimating μ_{p_i} when it is suspected *a priori* that $\mu_{p_i} = \mu_{p_o}$, in the form of the null hypothesis, $H_o : \mu_{p_i} = \mu_{p_o}$. The intent is to use this *UPI* to increase the efficiency of estimators by performing a preliminary test on the null hypothesis. This natural origin, μ_{p_o}, could be any type of prior knowledge about μ_{p_i}. In many applied problems, the experimenter has some guessed value of the parameter of interest, based on past experience from previous experiments.

As indicatated above, suppose that the population proportion of "successes" (cures, in terms of the efficacy of a drug, or deaths in terms of the susceptibility of organisms) for a given dosage is p_i. Suppose also that one samples from this population at the given dose level, and observes P_i successes, when drawing n_i independent trials. Then

$$P_i = p_i + u_i \tag{1.1}$$

where the u_i are independently distributed, each with the binomial distribution with mean zero and variance $p_i(1-p_i)/n_i$.

In studies of this kind, it is common to suppose that p_i is governed by the logistic distribution function,

$$p_i = [1 + \exp -(\gamma + \beta\mu_{p_i})]^{-1} \tag{1.2}$$

Thus, (1.1) and (1.2) can be used to obtain the odds ratio of success versus failure,

$$\frac{P_i}{1-P_i} = \frac{p_i}{1-p_i}\left(\frac{1+u_i/p_i}{1-u_i/[1-p_i]}\right) \tag{1.3}$$

or, the logarithm of the odds ratio of success versus failure,

$$\ln\left(\frac{P_i}{1-P_i}\right) = \ln\left(\frac{p_i}{1-p_i}\right) + \ln[1+u_i/p_i] - \ln[1-u_i/(1-p_i)] \tag{1.4}$$

Zellner and Lee (1965) motivate the estimation of such a model by first taking a Taylor series approximation to the last two terms, retaining linear terms and denoting the remainder, R_i, yielding the following relationship

$$\begin{aligned} \ln\left(\frac{P_i}{1-P_i}\right) &= \ln\left(\frac{p_i}{1-p_i}\right) + [u_i/p_i] + [u_i/(1-p_i)] + R_i \\ &= \gamma + \beta\mu_{p_i} + \{u_i/(p_i[1-p_i])\} + R_i \end{aligned} \tag{1.5}$$

Given the relationship, (1.2), and assuming that β is bounded away from from 0 with probability one, the logit of p_i, $logit(p_i) = \ln[p_i/(1-p_i)]$, can be expressed in terms of γ, β and μ_{p_i}, as follows

$$\gamma + \beta\mu_{p_i} = logit(p_i) = \ln\left(\frac{p_i}{1-p_i}\right), \quad 0 < p_i < 1. \tag{1.6}$$

Solving for μ_{p_i}, yields an expression for the effective dose with mortality rate, p_i,

$$\mu_{p_i} = \frac{1}{\beta}\left\{\ln\left(\frac{p_i}{1-p_i}\right) - \gamma\right\}. \tag{1.7}$$

Now, suppose that (γ, β) can be consistently estimated using the maximum likelihood estimator (*MLE*), $(\hat{\gamma}, \hat{\beta})$. Then, the *MLE* of μ_{p_i} can be obtained by replacing (γ, β) in (1.7) with $(\hat{\gamma}, \hat{\beta})$ yielding

$$\hat{\mu}_{p_i} = \frac{1}{\hat{\beta}}\left\{\ln\left(\frac{p_i}{1-p_i}\right) - \hat{\gamma}\right\}. \tag{1.8}$$

A frequently used value of $p_i = 0.5$, which we will denote simply as p in what follows. Then, we denote μ_{p_i} simply as μ_p, and this represents the median effective dose. This dose is commonly called ED_50, or LC_50. In this case $\mu_p = \gamma/\beta$ and its *MLE* is thus $\hat{\mu}_p = \hat{\gamma}/\hat{\beta}$. In general, as indicated above, $0 < p_i < 1$. Thus the discussion which follows regarding estimation can be applied to any level of dosage yielding p_i mortality without loss of genarality. This is because the random vector on which $\hat{\mu}_{p_i}$ for $p_i \neq 0.50$ depends is simply $(\hat{\gamma}, \hat{\beta})$.

The estimated dispersion matrix for $(\hat{\gamma}, \hat{\beta})'$ is

$$\hat{\mathbf{\Sigma}} = \begin{pmatrix} \hat{\sigma}_{11} & \hat{\sigma}_{12} \\ \hat{\sigma}_{21} & \hat{\sigma}_{22} \end{pmatrix}, \tag{1.9}$$

where $\hat{\sigma}_{11}$ denotes the estimated variance of $\hat{\gamma}$ and $\hat{\sigma}_{22}$ denotes the estimated variance of $\hat{\beta}$. The off diagonal elements of $\hat{\mathbf{\Sigma}}$ represent the estimated covariance between $\hat{\gamma}$ and $\hat{\beta}$. The elements of $\hat{\mathbf{\Sigma}}$ are used below to compute an estimator of $V[\hat{\mu}_p]$.

Suppose now that the *UPI* regarding μ_p is available, and that $\mu_p = \mu_{p_o}$. It is then reasonable to move the *MLE* of μ_p close to μ_{p_o}. This information can be used in various ways in constructing new and improved estimators of μ_p. First, we propose a *shrinkage estimator* (*SE*) of μ_p by combining sample and non-sample information as follows

$$\hat{\mu}_p^S = \pi\mu_{p_o} + (1-\pi)\hat{\mu}_p. \tag{1.10}$$

One might consider π $(0 < \pi \leq 1)$ and $(1-\pi)$ as two weighted constants, reflecting the prior relative confidence with which μ_{p_o} and $\hat{\mu}_p$ respectively are viewed. Thus, the choice of μ_{p_o} may be based on the experimenter's prior knowledge, and the choice of π represents the degree of trust the experimenter has in μ_{p_o}. Genrally speaking, the *SE* gives a smaller mean squared error at and near the origin, at the expense of poorer performance for large values of $|\mu_p - \mu_{p_o}|$. In an effort to avoid this undesirable feature of *SE*, and to increase the efficiency of

estimators when the prior information is rather uncertain, it may be profitable to construct a *shrinkage preliminary test estimator* (*SPTE*), denoted by $\hat{\mu}_p^{SP}$. In this case, the estimator $\hat{\mu}_p$ or $\hat{\mu}_p^S$ is selected according to whether $H_o : \mu_p = \mu_{p_o}$ is rejected or not. On the other hand, the usual preliminary test estimator (Bancroft (1944)) chooses the *MLE*, $\hat{\mu}_p$ or μ_{p_o} based on the outcome of the preliminary test. However, the usefulness of estimators based on the usual preliminary test principle (in terms of reductions in mean squared error) is limited depending on the size of the preliminary test. We show that the proposed estimation strategies significantly improve upon the performance of the usual preliminary test estimator for a given test size, in terms of mean squared error.

The remainder of the paper is structured as follows. In Section 2, we describe a series of estimation strategies for μ_p. Section 3 details the asymptotic performance of the various estimators, in terms of asymptotic bias and mean squared error. In Section 4, we compare the performance of the estimators based on their asymptotic mean squared errors. Section 5 outlines the asymptotic relative efficiency properties of the competing estimators. Section 6 provides analysis of various aspects of this type of estimation environment in terms of a series of simulation experiments. Section 7 summarises and conculdes.

2 Proposed Estimation Strategies

While *SE* depends on π, which in turn reflects the degree of confidence one has in the *UPI*, *SE* does *not* depend directly on any test statistic one might construct with respect to the null hypothesis, $H_o : \mu_p = \mu_{p_o}$. This is in contrast to the *SPTE*, defined in (1.10) above, which depends not only on such a test statistic but also on π. In addition, of course, the usual *preliminary test estimator* (*PTE*), depends on a test statistic of H_o, and the significance level of the test. Thus, in order to define *PTE* and a *shrinkage preliminary test estimator* (*SPTE*), one must first define an appropriate test statistic of H_o, embodying the *UPI*.

Since $\hat{\mu}_p = \hat{\gamma}/\hat{\beta}$, an estimator of $\mathrm{V}[\hat{\mu}_p]$ can be obtained using the elements of $\hat{\mathbf{\Sigma}}$. In particular, taking a first-order Taylor series approximation to $\hat{\gamma}/\hat{\beta}$ around μ_p, and ignoring higher-order terms yields

$$\frac{\hat{\gamma}}{\hat{\beta}} \approx \frac{\gamma}{\beta} + \frac{\partial[\hat{\gamma}/\hat{\beta}]}{\partial\hat{\gamma}}\bigg|_{(\gamma,\beta)}\{\hat{\gamma}-\gamma\} + \frac{\partial[\hat{\gamma}/\hat{\beta}]}{\partial\hat{\beta}}\bigg|_{(\gamma,\beta)}\{\hat{\beta}-\beta\} \tag{2.1}$$

Thus, $\mathrm{V}[\hat{\mu}_p]$ can be consistently estimated using the estimator, $\hat{\mathrm{V}}[\hat{\mu}_p]$, defined as follows,

$$\hat{\mathrm{V}}[\hat{\mu}_p] \equiv \hat{\sigma}^2 = \frac{1}{\hat{\beta}^2}(\hat{\sigma}_{11} + 2\hat{\mu}_p\hat{\sigma}_{12} + \hat{\mu}_p^2\hat{\sigma}_{22}) \tag{2.2}$$

which is the variance of the approximation on the right hand side of equation (2.1). Goedhart (1986) has demonstrated that the limiting distribution of $\sqrt{n}\,(\hat{\mu}_p - \mu_p) \xrightarrow{L} \mathcal{N}(0,\ \sigma^2)$, where $\xrightarrow{\mathcal{L}}$ means convergence in law (distribution) and $\sigma^2 = (1/\beta^2)(\sigma_{11} + 2\mu_p\sigma_{12} + \mu_p^2\sigma_{22})$.

In light of the foregoing, we propose the following test statistic for the preliminary test of H_o

$$D_n = \left\{\frac{\sqrt{n}(\hat{\mu}_p - \mu_{p_o})}{\hat{\sigma}}\right\}^2, \quad \hat{\sigma}^2 = \frac{1}{\hat{\beta}^2}(\hat{\sigma}_{11} + 2\hat{\mu}_p\hat{\sigma}_{12} + \hat{\mu}_p^2\hat{\sigma}_{22}). \tag{2.3}$$

Given the asymptotic distribution of $\hat{\mu}_p$, $D_n \overset{A}{\sim} \chi^2[1]$.

For a given level of significance, α (such that $0 < \alpha < 1$), let $d_\alpha = d_{n,\alpha}$ be the upper $100\alpha\%$ critical value using the distribution of D_n under H_o. Then, the *SPTE*, $\hat{\mu}_p^{SP}$, of μ_p can be constructed as follows,

$$\hat{\mu}_p^{SP} = \{(1-\pi)\hat{\mu}_p + \pi\mu_{p_o}\}I(D_n < d_\alpha) + \hat{\mu}_p I(D_n \geq d_\alpha), \tag{2.4}$$

where $I(A)$ is the indicator function of set A. For $\pi = 1$ in the above relation, we obtain the usual *PTE*

$$\hat{\mu}_p^{P} = \mu_{p_o} I(D_n < d_\alpha) + \hat{\mu}_p I(D_n \geq d_\alpha), \tag{2.5}$$

Thus $\hat{\mu}_p^P$ is a special case of $\hat{\mu}_p^{SP}$. It is important to note, however, that estimators based on a preliminary test rule do not uniformly outclass $\hat{\mu}_p$. More importantly, on the other hand, they do possess a bounded *mean squared error* (*MSE*) which is substantially smaller than the *MSE* of $\hat{\mu}_p$ in a reigon at and near the null hypothesis. On the other hand, $\hat{\mu}_p^S$ has an unbounded *MSE*.

It is also of interest to note that the statistic, D_n, is consistent against any fixed alternative, $\mu_p = \mu_{p_o}$. In this case, and for large n, $\hat{\mu}_p^{SP}, \hat{\mu}_p^P$ and $\hat{\mu}_p$ are all asymptotically equivalent. Consequently, we should specify a more interesting sequence of local alternatives to avoid such asymptotic degeneracy. Furthermore, Ahmed(1997), among others, has pointed out that estimators based on a preliminary test principle possess substantially smaller *MSE* than that of $\hat{\mu}_p$ in a shrinkage neighborhood of the *UPI* in the parameter space. For this reason, a sequence, $\{K_n\}$, of local alternatives defined as follows is considered:

$$K_n : \mu_{p_o} = \mu_{p_n} + \frac{\delta}{\sqrt{n}} \tag{2.6}$$

where δ is a fixed real number.

3 Asymptotic Results

To study the asymptotic properties of the proposed estimators, the following lemma is useful.

Lemma 3.1: Under the local alternatives in (2.6)

1. $\sqrt{n}(\hat{\mu}_p - \mu_{p_n}) \xrightarrow{\mathcal{L}} \mathcal{N}(\delta, \sigma^2)$, where $\sigma^2 = (1/\beta^2)(\sigma_{11} + 2\mu_p\sigma_{12} + \mu_p\sigma_{22})$

2. The test statistic D_n in (2.3) has an asymptotically noncentral chi-square distribution with 1 degree of freedom and noncentrality parameter, $\Delta = (\mu_p - \mu_{p_o})^2/\sigma^2$. Also, $d_{n,\alpha} \to \chi^2_\alpha[1]$ as $n \to \infty$. Thus, the critical value $d_{n,\alpha}$ of D_n may be approximated by $\chi^2_\alpha[1]$, the upper $100\alpha\%$ critical value of the central chi-square distribution with 1 degree of freedom.

Under the local alternatives, (2.6), and by virtue of Lemma 3.1, can derive expressions for the *asymptotic bias* (*AB*) of the estimators discussed above. Theorem 3.1 below states the *AB* of each of $\hat{\mu}_p, \hat{\mu}_p^S, \hat{\mu}_p^{SP}$ and $\hat{\mu}_p^P$.

Theorem 3.1:

$$AB(\hat{\mu}_p) = \lim_{n\to\infty} E\{\sqrt{n}(\hat{\mu}_p - \mu_{p_n})\} = 0 \tag{3.1}$$

$$AB(\hat{\mu}_p^S) = \lim_{n\to\infty} E\{\sqrt{n}(\hat{\mu}_p^S - \mu_{p_n})\} = \pi\delta, \quad \delta = \mu_p - \mu_{p_o} \tag{3.2}$$

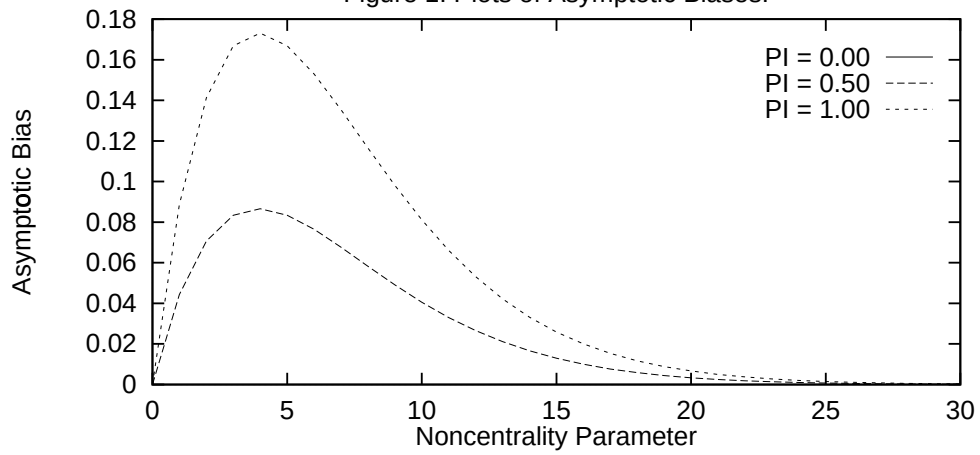
$$AB(\hat{\mu}_p^{SP}) = \lim_{n\to\infty} E\{\sqrt{n}(\hat{\mu}_p^{SP} - \mu_{p_n})\} = \pi\delta H_3(\chi^2_{1,\alpha};\Delta) \tag{3.3}$$

$$AB(\hat{\mu}_p^P) = \lim_{n\to\infty} E\{\sqrt{n}(\hat{\mu}_p^P - \mu_{p_n})\} = \delta H_3(\chi^2_{1,\alpha};\Delta) \tag{3.4}$$

where $H_v(.;\Delta)$ is the cumulative distribution of a noncentral chi-square distribution with v degrees of freedom and noncentrality parameter, Δ.

It can be seem from Theorem 3.1 that $AB(\hat{\mu}_p^{SP}) \leq AB(\hat{\mu}_p^P)$, since $\pi \in (0,1]$, and we conclude that $\hat{\mu}_p^{SP}$ has less asymptotic bias than $\hat{\mu}_p^{SP}$, depending upon the value of π. The $AB(\hat{\mu}_p^{SP})$ and $AB(\hat{\mu}_p^P)$ are zero when $\Delta = 0$, increase to maxima, then decrease towards zero as Δ increases. Figure 1 illustrates the behaviour of the AB functions for $\hat{\mu}_p$ (where $\pi = 0$), $\hat{\mu}_p^{SP}$ (with $\pi = 0.5$) and $\hat{\mu}_p^P$ (which is equivalent to $\hat{\mu}_p^{SP}$ with $\pi = 1$).

Figure 1: Plots of Asymptotic Biases.



It is thus evident from Figure 1 that the AB function of $\hat{\mu}_p^P$ rises more quickly than that of $\hat{\mu}_p^{SP}$, and reaches a higher maximum. Furthermore, as Δ rises, $\hat{\mu}_p^{SP}$ approaches zero more quickly than $\hat{\mu}_p^P$. This is also indicated by the fact that the standard deviation of $AB(\hat{\mu}_p^{SP})$ as a function of Δ is approximately half that of the standard deviation of $AB(\hat{\mu}_p^P)$.

Since bias, in general, is part of *MSE*, and control of *MSE* will control both bias and variance, we shall focus from this point onwards on *MSE*. Furthermore, as the analysis is being conducted in an asymptotic framework, we will define and discuss *asymptotic mean squared error* (*AMSE*).

Under the local alternatives (2.6) and by virtue of Lemma 3.1, we provide the expressions for the *AMSE* of the estimators under study in the following theorem.

Theorem 3.2: The *AMSE* for $\hat{\mu}_p, \hat{\mu}_p^S, \hat{\mu}_p^{SP}$ and $\hat{\mu}_p^P$ are as specified in equations (3.4)–(3.7) below, respectively:

$$AMSE(\hat{\mu}_p) = \sigma^2 \tag{3.4}$$

$$AMSE(\hat{\mu}_p^S) = \sigma^2(1-\pi)^2 + \sigma^2\Delta\pi^2 \tag{3.5}$$

$$AMSE(\hat{\mu}_p^{SP}) = \sigma^2 - \sigma^2\pi(2-\pi)H_3(\chi^2_\alpha[1];\Delta) + \sigma^2\Delta\{2\pi H_3(\chi^2_\alpha[1];\Delta) - \pi(2-\pi))H_5(\chi^2_\alpha[1];\Delta)\}. \tag{3.6}$$

$$AMSE(\hat{\mu}_p^{P}) = \sigma^2 - \sigma^2 H_3(\chi^2_\alpha[1];\Delta) + \sigma^2\Delta\{2H_3(\chi^2_\alpha[1];\Delta) - H_5(\chi^2_\alpha[1];\Delta)\}. \tag{3.7}$$

Proof: The foregoing expressions for $AMSE(\hat{\mu}_p^{SP})$ and $AMSE(\hat{\mu}_p^{P})$ can be obtained using analogous manipulations to those presented in Appendix B of Judge and Bock (1978).

4 Comparisons of *AMSE* for $\hat{\mu}_p, \hat{\mu}_p^S, \hat{\mu}_p^{SP}$ and $\hat{\mu}_p^P$

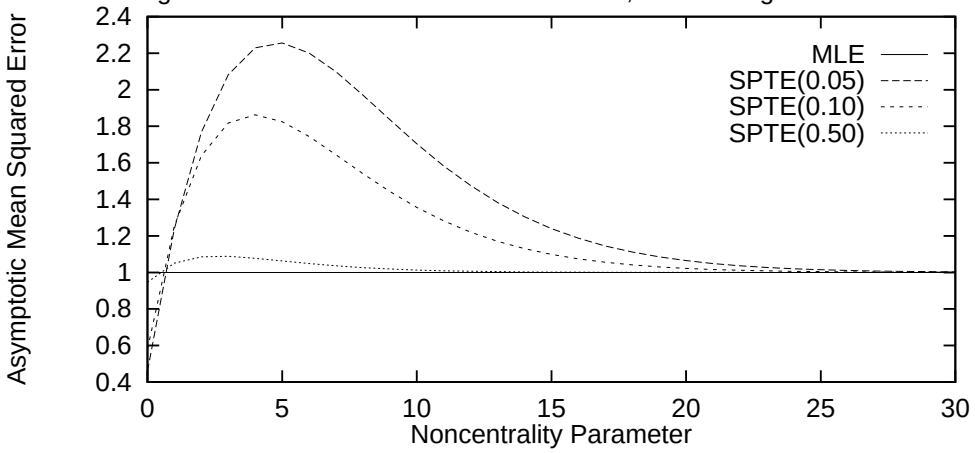
4.1 Comparing the MLE ($\hat{\mu}_p$) with the SE ($\hat{\mu}_p^S$)

Theorem 3.2 indicates that the *AMSE* of $\hat{\mu}_p$ is constant for all Δ, while the *AMSE* of $\hat{\mu}_p^S$ is a straight line emanating from the point $\sigma^2(1-\pi)^2$, with a slope of $\sigma^2\pi^2$. $AMSE(\hat{\mu}_p)$ is intersected by $AMSE(\hat{\mu}_p^S)$ when $\Delta = (2-\pi)/\pi$. Thus, $AMSE(\hat{\mu}_p^S)$ is less than $AMSE(\hat{\mu}_p)$ when $\Delta \in [0, (2-\pi)/\pi)$. Hence, for Δ at and close to zero, $\hat{\mu}_p^S$ performs better than $\hat{\mu}_p$ in terms of *AMSE*. Alternatively, as Δ deviates from the origin, $AMSE(\hat{\mu}_p^S)$ increases without bound. Thus, departures from the restriction, $\mu_p = \mu_{p_o}$ is fatal to $\hat{\mu}_p^S$ but is of no concern to $\hat{\mu}_p$. That is, $\hat{\mu}_p$ has constant *AMSE* for all Δ.

4.2 Comparing the MLE ($\hat{\mu}_p$) with the SPTE ($\hat{\mu}_p^{SP}$)

Turning to a consideration of *AMSE* performances of $\hat{\mu}_p$ and $\hat{\mu}_p^{SP}$, it is clear that for Δ equal to or close to zero, $AMSE(\hat{\mu}_p^{SP})$ is less than $AMSE(\hat{\mu}_p)$. As α, the level of statistical significance associated with a test of $H_0 : \mu_p = \mu_{p_o}$, approaches one, $AMSE(\hat{\mu}_p^{SP})$ tends to $AMSE(\hat{\mu}_p)$. Also, when Δ increases and tends to infinity, $AMSE(\hat{\mu}_p^{SP})$ approaches $AMSE(\hat{\mu}_p)$. More generally, as Δ rises (given α), the value of $AMSE(\hat{\mu}_p^{SP})$ increases, crosses $AMSE(\hat{\mu}_p)$, reaches a maximum, and then monotonically decreases, approaching $AMSE(\hat{\mu}_p)$ as $\Delta \to \infty$. Figure 2 shows the behaviour of the *AMSE* functions for three values of $\alpha = \{0.05, 0.10, 0.50\}$, given $\pi = 0.5$. For convenience, and without loss of generality, Figure 2 is drawn with σ^2 normalised at unity.

Figure 2: Plot of AMSE for MLE and SPTE, Various Significance Levels.



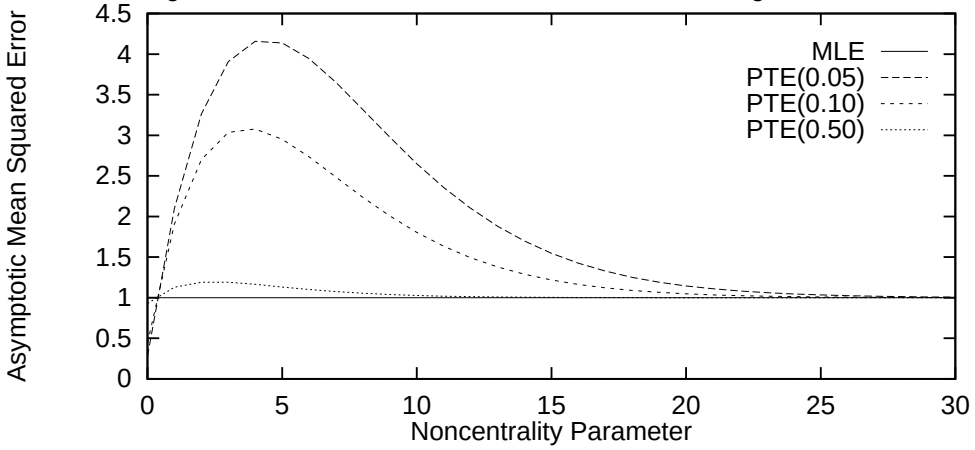
4.3 Comparing the MLE ($\hat{\mu}_p$) with the PTE ($\hat{\mu}_p^P$)

By manipulation of (3.7) it can be shown that, as $\alpha \to 0$, $AMSE(\hat{\mu}_p^P)$ will be less than $AMSE(\hat{\mu}_p)$, so long as $0 \leq \Delta \leq 1$. In particular, we know that $AMSE(\hat{\mu}_p) = \sigma^2$ from equation (3.4). Setting this equal to (3.7) and solving for Δ ($\Delta_{1,4}$) at which this equality holds yields

$$\Delta_{1,4} = \left(\frac{H_3(\chi^2_\alpha[1];\Delta)}{2H_3(\chi^2_\alpha[1];\Delta) - H_5(\chi^2_\alpha[1];\Delta)} \right) \tag{4.1}$$

Then, as $\alpha \to 0$, $H_3(\chi^2_\alpha[1];\Delta) \to 1$ and $H_5(\chi^2_\alpha[1];\Delta) \to 1$, so that $\lim_{\alpha \to 0} \Delta_{1,4} = 1$. However, for $\alpha > 0$, $\Delta_{1,4} < 1$. Also, since $AMSE(\hat{\mu}_p^S)$ cuts $AMSE(\hat{\mu}_p)$ at $(2-\pi)/\pi > 1$, the range of Δ for which $AMSE(\hat{\mu}_p^P)$ is less than $AMSE(\hat{\mu}_p)$ is less than the range of Δ for which $AMSE(\hat{\mu}_p^S)$ is less than $AMSE(\hat{\mu}_p)$. Figure 3 shows the behaviour of $AMSE(\hat{\mu}_p^P)$ functions for three values of $\alpha = \{0.05, 0.10, 0.50\}$, for $\pi = 0.5$. Again, as with Figure 2, Figure 3 is drawn with σ^2 normalised at unity. In Figure 3, it can be seen that $AMSE(\hat{\mu}_p^P)$ attains higher values at its maxima for given values of α than $AMSE(\hat{\mu}_p^{SP})$. This will be discussed in more detail in the next Sub-section.

Figure 3: Plot of AMSE for MLE and PTE, Various Significance Levels.



4.4 Comparing the SPTE ($\hat{\mu}_p^{SP}$) with the PTE ($\hat{\mu}_p^P$)

Equations (3.4) and (3.6) can be used to determine the range of values of Δ for which $AMSE(\hat{\mu}_p^{SP})$ is less than $AMSE(\hat{\mu}_p)$ in the same way as $\Delta_{1,4}$ was identified earlier. This range of Δ can then be compared with $\Delta \in [0, \Delta_{1,4}]$. Using (3.4) and (3.6), we solve for Δ ($\Delta_{1,3}$) at which $AMSE(\hat{\mu}_p) = AMSE(\hat{\mu}_p^{SP})$,

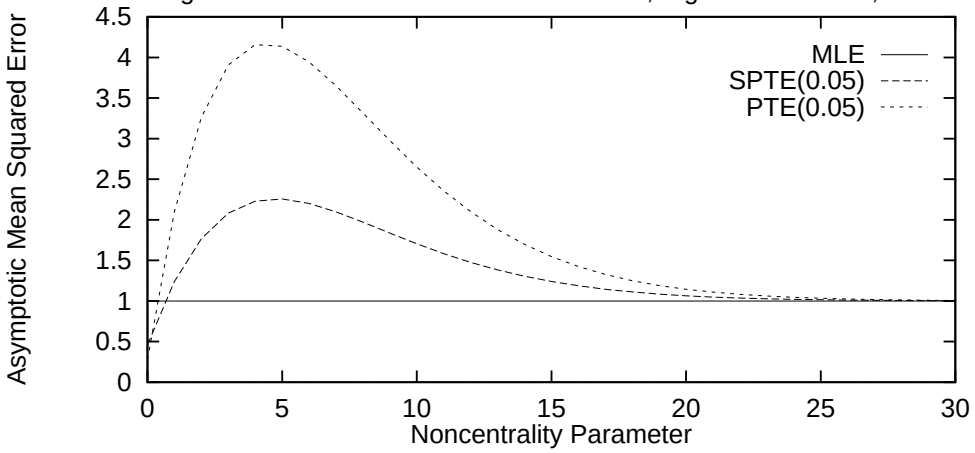
$$\Delta_{1,3} = \left(\frac{2-\pi}{\pi} \right) \left(\frac{\pi H_3(\chi^2_\alpha[1];\Delta)}{2H_3(\chi^2_\alpha[1];\Delta) - (2-\pi)H_5(\chi^2_\alpha[1];\Delta)} \right) \tag{4.2}$$

Again, as $\alpha \to 0$, $H_3(\chi^2_\alpha[1];\Delta) \to 1$ and $H_5(\chi^2_\alpha[1];\Delta) \to 1$, so that $\lim_{\alpha \to 0} \Delta_{1,3} = (2-\pi)/\pi$. However, for $\alpha > 0$, $\Delta_{1,3} > (2-\pi)/\pi$, since the expression in the second parentheses on the right hand side of (4.2),

$$\left(\frac{\pi H_3(\chi^2_\alpha[1];\Delta)}{2H_3(\chi^2_\alpha[1];\Delta) - (2-\pi)H_5(\chi^2_\alpha[1];\Delta)} \right) \tag{4.3}$$

is greater than one. This can be seen by taking limits of the numerator and denominator of (4.3) as $\pi \to 0$. The further implication of this last point is that, for $\alpha \neq 0$, $\Delta \in [0, \Delta_{1,3}]$ is *wider* than $\Delta \in [0, (2-\pi)/\pi]$, the range of Δ for which $AMSE(\hat{\mu}_p^S) < AMSE(\hat{\mu}_p)$. Figure 4 below shows the plots of $AMSE(\hat{\mu}_p^{SP})$ (for $\pi = 0.5$) and $AMSE(\hat{\mu}_p^P)$ at $\alpha = 0.05$.

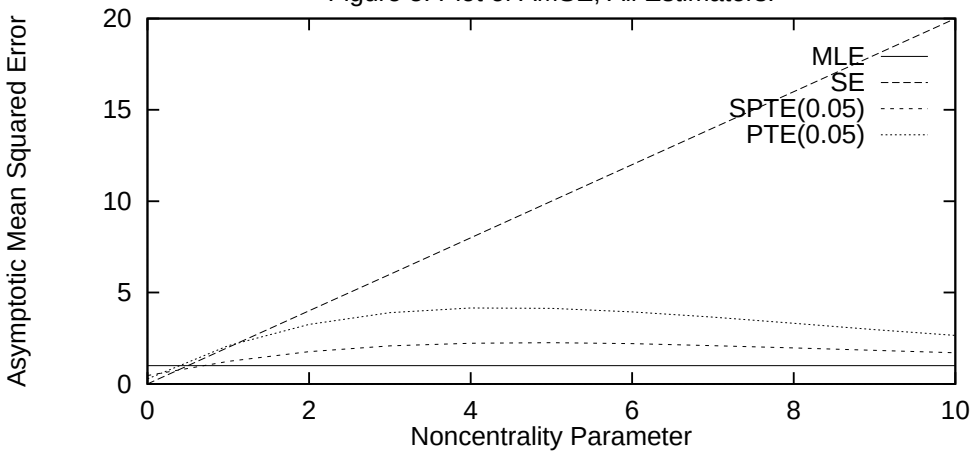
Figure 4: Plot of AMSE for SPTE and PTE, Significance Level, 0.05.



4.5 Comparing AMSE Functions for All Estimators

A useful summary of the foregoing discussion can be provided by plotting representative *AMSE* functions for each of the estimators, $\hat{\mu}_p, \hat{\mu}_p^S, \hat{\mu}_p^{SP}$ and $\hat{\mu}_p^P$. This is presented in Figure 5, for $\pi = 0.5$ where applicable, and $\alpha = 0.05$.

Figure 5: Plot of AMSE, All Estimators.



The discussion of the preceding Sub-sections and the plots of the *AMSE* functions in Figure 5 above permits a more formal statement of the relative dominance characteristics of the various estimators. Recall that when $\Delta \in [0, \Delta_{1,3}]$, we have $AMSE(\hat{\mu}_p^{SP}) \leq AMSE(\hat{\mu}_p)$, which is a wider interval than $\Delta \in [0, \Delta_{1,4}]$. Thus, $\hat{\mu}_p^{SP}$ provides a wider range of Δ than $\hat{\mu}_p^P$, the usual *PTE*, for which $\hat{\mu}_p^{SP}$ dominates $\hat{\mu}_p$. This indicates the superiority of $\hat{\mu}_p^{SP}$ over $\hat{\mu}_p^P$. Furthermore, equations (3.6) and (3.7) indicate that $AMSE(\hat{\mu}_p^P) - AMSE(\hat{\mu}_p^{SP})$ is given by

$$\sigma^2\Delta\left\{2(\,1-\pi)H_3(\chi^2_\alpha[1];\Delta) - (1-\pi)^2 H_5(\chi^2_\alpha[1];\Delta)\right\} - \sigma^2(1-\pi)^2 H_3(\chi^2_\alpha[1];\Delta). \qquad (4.4)$$

The magnitude of (4.4) can be positive or negative, depending on the magnitude of Δ. That is, for Δ in a neighbourhood of zero, (4.4) is negative, since $-\sigma^2(1-\pi)^2 H_3(\chi^2_\alpha[1];\Delta) < 0$, and $\sigma^2\Delta\{2(1-\pi)H_3(\chi^2_\alpha[1];\Delta) - (1-\pi)^2 H_5(\chi^2_\alpha[1];\Delta)\}$ is positive but negligible, especially for larger values of π. Thus, when the former term is larger in absolute value than the latter, $AMSE(\hat{\mu}_p^P) \leq AMSE(\hat{\mu}_p^{SP})$. For such small values of Δ, $\hat{\mu}_p^P$ dominates $\hat{\mu}_p^{SP}$. However, as Δ rises, the difference $AMSE(\hat{\mu}_p^P) - AMSE(\hat{\mu}_p^{SP})$ becomes positive, so that $\hat{\mu}_p^P$ is dominated uniformly by $\hat{\mu}_p^{SP}$ for such larger values of Δ, in the rest of the parameter space.

More formally, suppose Δ_π is a point in the parameter space at which the $AMSE(\hat{\mu}_p^{SP}) = AMSE(\hat{\mu}_p^P)$ for a given value of π. Then, for $\Delta \in [0, \Delta_\pi]$, $\hat{\mu}_p^P$ performs better than $\hat{\mu}_p^{SP}$, while for $\Delta \in (\Delta_\pi, \infty)$, $\hat{\mu}_p^{SP}$ dominates $\hat{\mu}_p^P$ uniformly. Further, for larger values of π (close to unity), the interval $[0, \Delta_\pi]$ may be not noticeable. Nevertheless, $\hat{\mu}_p^{SP}$ and $\hat{\mu}_p^P$ share the common asymptotic property that their $AMSE$ converge to a common limit, $AMSE(\hat{\mu}_p) = \sigma^2$, as $\Delta \to \infty$,

Figures 2 and 3 above also indicate the inherent variability of $AMSE(\hat{\mu}_p^{SP})$ and $AMSE(\hat{\mu}_p^P)$ (at $\pi = 0.5$), as functions of α. Each of the functions exhibits increasing variability as α falls. In addition, for a given α (again conditional on $\pi = 0.5$), $AMSE(\hat{\mu}_p^P)$ exhibits greater variability than $AMSE(\hat{\mu}_p^{SP})$. The standard deviation of $AMSE(\hat{\mu}_p^{SP})$ is half that of $AMSE(\hat{\mu}_p^P)$.

We conclude this section with the following proposition, which suggests that relative *percentage improvement* (PI) statistics can be computed for each of the estimators in terms of their $AMSE$, relative to the $AMSE(\hat{\mu}_p)$.

Proposition 4.1: None of the four estimators is inadmissible with respect to the other three. However, at $\Delta = 0$, the estimators may be ordered according to the magnitude of their $AMSE$ as follows:

$$\hat{\mu}_p^S \succ \hat{\mu}_p^P \succ \hat{\mu}_p^{SP} \succ \hat{\mu}_p, \tag{4.5}$$

where $\succ$ denotes domination.

To appraise the estimators at $\Delta = 0$, we compute the PI in $AMSE$ as follows:

$$PI_{\hat{\mu}_p^i} = \frac{100(AMSE(\hat{\mu}_p) - AMSE(\mu_p^i)}{AMSE(\hat{\mu}_p)} \quad i = S, SP, P \tag{4.6}$$

Table 1 below gives these PI at $\Delta = 0$. Table 1 compares $PI_{\hat{\mu}_p^S}$ to $PI_{\hat{\mu}_p^{SP}}$ at three values of π when $\Delta = 0$, in the first three columns of the Table. The values of $PI_{\hat{\mu}_p^{SP}}$ are conditional on $\alpha = 0.05$. As one would expect, the PI increase as π rises (meaning greater weight being placed on the restriction that $\mu_p = \mu_{p_o}$). This being true, of course, in light of the restriction being true, since $\Delta = 0$. The second series of three columns in Table 1 compare $PI_{\hat{\mu}_p^{SP}}$ to $PI_{\hat{\mu}_p^P}$ at $\pi = 0.5$, for three values of α. Clearly, $PI_{\hat{\mu}_p^P}$ is higher than $PI_{\hat{\mu}_p^{SP}}$, since $\mu_p = \mu_{p_o}$ holds. However, for both types of comparison in Table 1, the superiority of $\hat{\mu}_p^S$ and $\hat{\mu}_p^P$ over $\hat{\mu}_p^{SP}$ is small enough to justify the use of $\hat{\mu}_p^{SP}$, given its dominance characteristics when Δ deviates from zero.

Table 1: Percentage Improvements
(PI) Over $\hat{\mu}_p$ at $\Delta = 0$

$\alpha = 0.05$			$\pi = 0.50$		
π	$PI_{\hat{\mu}_p^S}$ %	$PI_{\hat{\mu}_p^{SP}}$ %	α	$PI_{\hat{\mu}_p^{SP}}$ %	$PI_{\hat{\mu}_p^P}$ %
0.25	43.75	31.54	0.05	54.07	72.09
0.50	75.00	54.07	0.10	42.05	56.07
0.75	93.75	67.59	0.25	20.73	27.64

In general, whether $\mu_p = \mu_{p_o}$ is not known. Also, $AMSE(\hat{\mu}_p^{SP})$ and $PI_{\hat{\mu}_p^{SP}}$ depend on α, which must be determined by the researcher. One method to determine α is to compute the minimum guaranteed efficiency of $\hat{\mu}_p^{SP}$. To do so, we turn to a consideration of *asymptotic relative efficiency*, which is defined and discussed in the following section.

5 Asymptotic Efficiency Analysis

The *asymptotic relative efficiency*, ARE, of an estimator is defined relative to $AMSE(\hat{\mu}_p)$, as follows

$$ARE(\hat{\mu}_p^i) = \frac{AMSE(\hat{\mu}_p)}{AMSE(\hat{\mu}_p^i)} \quad i = S, SP, P \tag{5.1}$$

¿From this definition, it is evident that ARE greater than unity signifies improvement of the estimator in question over $\hat{\mu}_p$. Section 4 indicates that $\hat{\mu}_p^{SP}$ is a preferred estimator in many situations, relative to $\hat{\mu}_p^S$ and $\hat{\mu}_p^P$. Thus, in what follows, we confine attention to $ARE(\hat{\mu}_p^{SP})$, which is a function of (α, π, Δ). For any $\alpha \neq 0$, $ARE(\hat{\mu}_p^{SP})$ has its maximum, (E^*), at $\Delta = 0$. Using equations (3.4) and (3.6), we can solve for E^* as follows

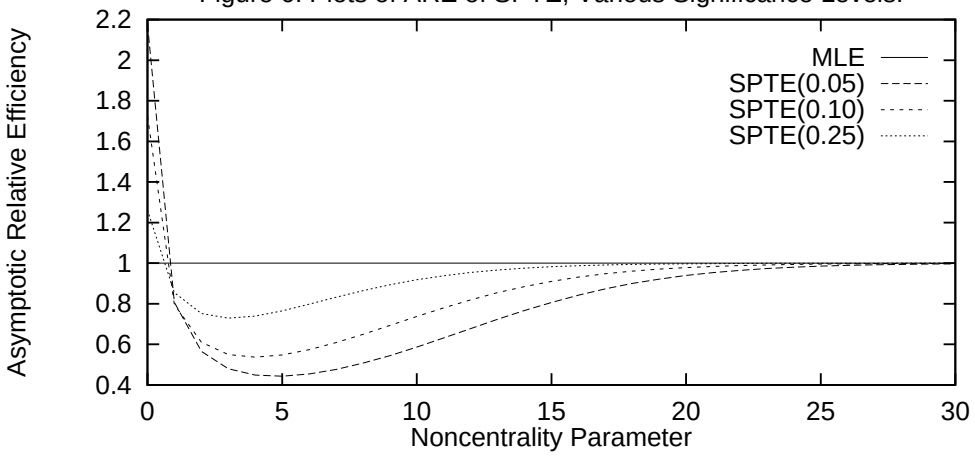
$$\begin{aligned} E^* =& \{1 - (2 - \pi)\pi H_3(\chi^2_\alpha[1]; 0)\}^{-1} \\ =& \{1 - (1 - (1 - \pi)^2) H_3(\chi^2_\alpha[1]; 0)\}^{-1}, \quad (\geq 1) \end{aligned} \tag{5.2}$$

For fixed values of α and π, $ARE(\hat{\mu}_p^{SP})$ decreases as Δ increases from zero, crossing the line $ARE(\hat{\mu}_p) = AMSE(\hat{\mu}_p)/AMSE(\hat{\mu}_p) = 1$ at $\Delta_{1,3}$ (the Δ consistent with solution of equation (4.2)), attains a minimum value at a point, Δ_o, then increases and asymptotically approaches $ARE(\hat{\mu}_p) = 1$.

It can also be noted that, for fixed π maximum efficiency, E^*, is a *decreasing* function of α. Minimum efficiency, E_o, is an *increasing* function of α, however. These features of $ARE(\hat{\mu}_p^{SP})$ are illustrated in Figure 6, which plots $ARE(\hat{\mu}_p^{SP})$ for $\alpha = (0.05,\ 0.10,\ 0.25)$, at

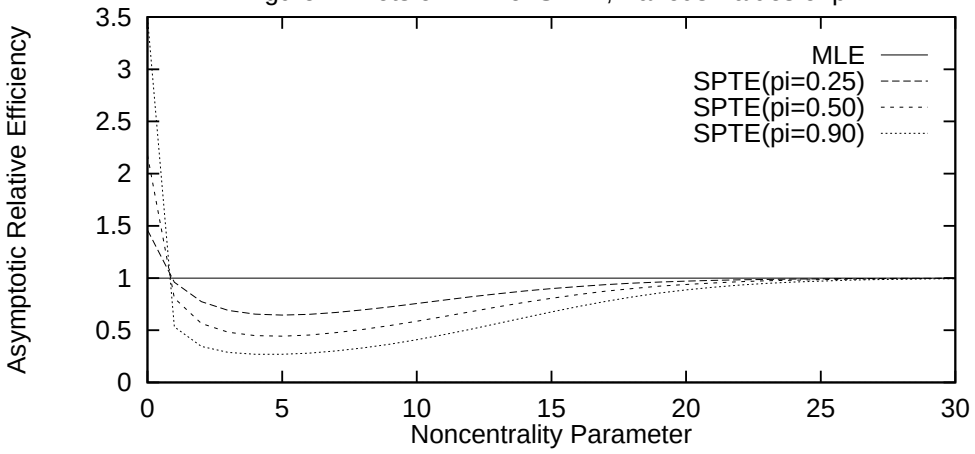
$\pi = 0.50$. It is also the case that, the *smaller* the value of α, the *greater* is the variability in $ARE(\hat{\mu}_p^{SP})$. The standard deviation of $ARE(\hat{\mu}_p^{SP})$ falls from 0.26 at $\alpha = 0.05$, to 0.09 at $\alpha = 0.25$.

Figure 6: Plots of ARE of SPTE, Various Significance Levels.



On the other hand, for any fixed value of α, the maximum efficiency of $ARE(\hat{\mu}_p^{SP})$, E^* is an *increasing* function of π, while the minimum efficiency of $ARE(\hat{\mu}_p^{SP})$, E_o, is a *decreasing* function of π. These features of $ARE(\hat{\mu}_p^{SP})$ are illustrated in Figure 7, which plots $ARE(\hat{\mu}_p^{SP})$ for $\pi = (0.25,\ 0.50,\ 0.90)$, at $\alpha = 0.05$. In this case, the *larger* the value of π, the *smaller* is the variability in $ARE(\hat{\mu}_p^{SP})$. The standard deviation of $ARE(\hat{\mu}_p^{SP})$ rises from 0.13 at $\pi = 0.25$, to 0.43 at $\pi = 0.90$.

Figure 7: Plots of ARE of SPTE, Various Values of pi.



In an effort to help the researcher in choosing an estimator with maximum ARE, we adopt the following procedure. If $\Delta \leq (2-\pi)/\pi$, then $\hat{\mu}_p^S$ is chosen because the performance of $\hat{\mu}_p^S$ is superior in this range. Since, however, Δ and π are generally unknown, there is no way of choosing a uniformly best estimator. We therefore recommend following a two step procedure for selecting the size of the preliminary test:

Step 1. Suppose the experimenter does not know the size of the test, but knows $\pi = \pi_o$, and wishes to choose an estimator, $\hat{\mu}_p^{SP}$, which has ARE not less than E_o. Then the `max-min` principle determines $\alpha = \alpha^*$ such that $ARE(\hat{\mu}_p^{SP}[\alpha^*, \pi_o, \Delta_o]) = E_o$. Therefore, a user who

wishes to find a good alternative to $\hat{\mu}_p^S$ and $\hat{\mu}_p$ should be able to specify the minimum ARE (E_o) to achieve this goal.

Step 2. Suppose the experimenter does not know the value of α *or* π, and wants to use the $\hat{\mu}_p^{SP}$ which has ARE no less than E_o. Then, among the set of estimators with $\alpha \in \mathcal{Z}$, where $\mathcal{Z} = \{\alpha : ARE(\hat{\mu}_p^{SP}) \geq E_o \text{ for all } \Delta \text{ and } \pi\}$, the estimator, $\hat{\mu}_p^{SP}$, is chosen to maximize the $ARE(\hat{\mu}_p^{SP})$ over all $\alpha \in \mathcal{Z}$ and all Δ, π. That is to say, we solve the following equation for α:

$$\max_{\alpha \in Z} \{\min_{\Delta} ARE(\hat{\mu}_p^{SP}[\alpha^*, \pi, \Delta_o]) = E_o\}, \quad \text{for all } \pi. \tag{5.3}$$

The solution, α^*, is the optimum level of significance.

Suppose, $\pi \in \Gamma$, where

$$\Gamma = \{\pi : \pi = 0(0.1)1\} \tag{5.4}$$

and the experimenter does not know the value of π but is willing to use an estimator with relative efficiency no less than E_o. Then $\alpha = \alpha^*$ can be found as described above, and the value of π can then also be determined.

As a result of the procedures in *Step 1* and *Step 2*, a researcher can select an estimator, $\hat{\mu}_p^{SP}$, with $ARE(\hat{\mu}_p^{SP}) > 1$. An example serves to illustrate the use of the procedure.

Example: If $\pi = 0.4$ and the researcher wishes to select an estimator, $\hat{\mu}_p^{SP}$, with a minimum ARE of at least 0.60, then the recommended $\alpha^* = 0.10$. This choice of α^* yields an estimator with a maximum ARE of 1.560. On the other hand, if the researcher wishes to rely on the *UPI* completely and uses $\pi = 1$, then the size of the preliminary test will be approximately 0.30. Also, the maximum ARE drops from 1.560 to 1.277.

The above example serves to illustrate the inherent usefulness of $\hat{\mu}_p^{SP}$ over $\hat{\mu}_p^P$. That is, for relatively larger values of Δ, the ARE of $\hat{\mu}_p^P$ will be smaller than that of $\hat{\mu}_p^{SP}$ for a given significance level. This can be inferred from the comparison of $AMSE(\hat{\mu}_p^{SP})$ and $AMSE(\hat{\mu}_p^P)$ in Figure 4, which is the underlying factor driving the performance of ARE in general. A table giving E^* and E_o fo a range of values of α and π is available from the authors in an appendix.

6 Simulation Experiments

In this section, we utilise the theoretical methodology of the preceding sections to determine the practical performance of the $\hat{\mu}_p^{SP}$ and $\hat{\mu}_p^P$, relative to $\hat{\mu}_p$ in a simulated setting. This setting is intended to replicate the circumstances an applied researcher is likely to face in the estimation of ED_{50}.

The baseline for the simulations conducted was the *MLE* simulation experiment design in Hamilton (1979). In that design, it is supposed that p_i follows the logistic distribution, as indicated in (1.2). In this context, μ_{p_i} is regarded as a dosage level, and is denoted as x_i in Hamilton (1979), who considers ten "dosage levels", $\mu_{p_i} = x_i = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$. Given these design values for μ_{p_i}, the mean value around which these are distributed is

Table 2: Dose Levels, $\mu_{p_i} = x_i$,
and Probabilities of Response, $p(x_i)$.

Dose x_i	Probability $p(x_i)$	Hamilton (1979) $p(x_i)$
$x_i = 1.0$	0.0001233946	0.00026
$x_i = 2.0$	0.0009110512	0.00160
$x_i = 3.0$	0.0066928510	0.01000
$x_i = 4.0$	0.0474258700	0.05969
$x_i = 5.0$	0.2689414000	0.28516
$x_i = 5.5$	0.50	0.50
$x_i = 6.0$	0.7310586000	0.71484
$x_i = 7.0$	0.9525741000	0.94031
$x_i = 8.0$	0.9933071000	0.99000
$x_i = 9.0$	0.9990889000	0.99840
$x_i = 10.0$	0.9998766000	0.99974

$\mu_{p_0} = 5.5$. Table 4, column I, of Hamilton (1979) indicates the p_i, which we can also denote as $p(x_i)$, for consistency with his notation. These probabilities are reproduced in Table 2. Using $p(x_i)$ in Table 2, it is possible to solve for γ and β, the parameters of the logistic distribution on which (1.2) is to be based for the simulations. This yields values of $\gamma = -10.088917$, and $\beta = 1.8343476$. For convenience in the simulation experiments in this paper, we therefore set $\gamma = -11.0$, and $\beta = 2.0$. Thus, $\mu_{p_0} = -(\gamma/\beta) = 5.50$, as is required for the ED_{50} dosage level. The second column in Table 2 thus indicates the $p(x_i)$ which are to be used in a series of four simulation experiments, and it can be seen that these are close to Hamilton's (1979) values for $p(x_i)$.

Including the value for $\mu_{p_0} = 5.5$, and the ten values in $x_i = \{1, \ldots, 10\}$, yields a total of eleven "observations" or dosage levels for one of the simulation experiments. Clearly, since one has to estimate γ and β in order to estimate μ_{p_i}, estimation based on such a régime would yield 11-2 = 9 degrees of freedom. Thus, estimators of γ and β would be subject to small-sample imprecision. In order to determine the effect of this on estimation of ED_{50}, we therefore conduct three other experiments based on this initial design. These additional experiments specify a finer grid for x_i, yielding 21, 50 and 105 dosage levels on these increasingly finer grids. Given the values for $\gamma = -11.0$ and $\beta = 2.0$, the various $p(x_i)$ can be obtained for subsequent generation of the required data for the simulation experiments. A full summary of x_i and $p(x_i)$ for these additional experiments is available from the authors in an appendix.

To construct the data required for the simulation experiments, first P_i denoted in equa-

tion (1.1) must be generated. The p_{x_i} for each x_i, and the respective experiments were therefore used to generate P_i, where P_i is the sample proportion of successes from a binomial distribution with population proportion of success of p_{x_i}, from n_i independent trials. The *IMSL* subroutine *RNBIN* was used to generate these success rates, based on $n_i = 500,000$ at each p_{x_i}. This ensured that, even for relatively small values of p_{x_i}, the success rate was non-zero. Thus, sample proportions could always be calculated. A total of 5,000 replicates were generated for each experiment (sample sizes of $N = \{11, 21, 50, 105\}$. In what follows, samples will be denoted $N_j, j = 1, \ldots, 4$. In addition, these 5,000 replicates were repeated 100 times, to yield a total of 500,000 samples. Thus, the performance of $\hat{\mu}_p^{SP}$ and $\hat{\mu}_p^{P}$ relative to $\hat{\mu}_p$ could be compared as Δ deviated from zero, at up to ninety-nine additional data points, based on 5,000 replicates each.

Having obtained P_i, u_i for equation (1.5) can be obtained as $P_i - p_{x_i}$. To complete data generation, R_i for equation (1.5) was generated as $R_i \sim (0, \sigma_R^2)$, where $\sigma_R^2 = 0.25$. The *IMSL* subroutine *DRNNOA* was used for this purpose. Given x_i, u_i, R_i, γ and β, the left-hand side of (1.5), the sample *logit*(P_i), can be calculated, from which one can then estimate γ and β. These estimates were obtained using the generalised least squares (GLS) routine in *SHAZAM*, Version 8.0 (White, 1978), adjusted to take account of the inherent heteroskedasticity in the disturbances of the model, (1.5).

In order to obtain $\hat{\mu}_p^{SP}$ and $\hat{\mu}_p^{P}$, the sample counterpart of D_n, D, must be computed for each replication, and compared with d_α. To control for possible experimental and small sample errors, especially when $N_j = (N_1, N_2) = (11, 21)$, the sample critical values of D are used. As indicated in Lemma 3.1, $d_{n,\alpha} \to \chi_\alpha^2[1]$ as $n \to \infty$. However, the relatively poor small sample performance of D can be seen by comparison of sample critical values of D at the various sample sizes, with the asymptotic critical values given a variety of α levels. These comparisons are given in Table 3.

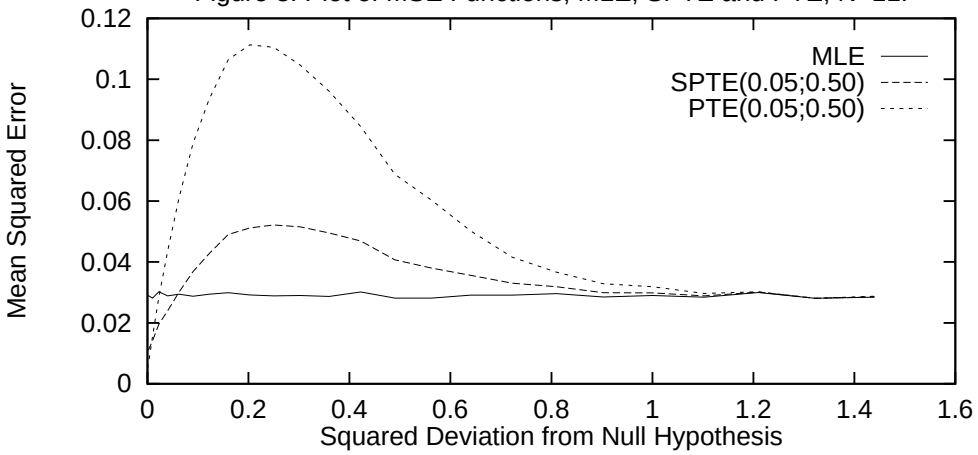
The sample critical values are obtained by using the 500,000 sample observations on D for each of the four simulation experiments with differing sample sizes. The values of D are sorted from lowest to highest, and the $100\alpha\%$ value obtained from the sorted vectors of D. It is clear that sample values of d_α approach the asymptotically valid values as N_j rises. However, even at $N_4 = 105$, there is still a substantial difference between the sample and asymptotic values. This indicates that one should exercise caution in practice in the use of the asymptotic critical values when computing $\hat{\mu}_p^{SP}$ or $\hat{\mu}_p^{P}$.

The next part of this Section deals with comparisons of the sample *MSE* functions of $\hat{\mu}_p^{SP}$ and $\hat{\mu}_p^{P}$ with $\hat{\mu}_p$. Figure 8 provides such a comparison for $N_1 = 11$, with $\pi = 0.5$ for $\hat{\mu}_p^{SP}$, and at an α level of 0.05.

Table 3: Sample and Asymptotic Critical Values for D and D_n, Various α.

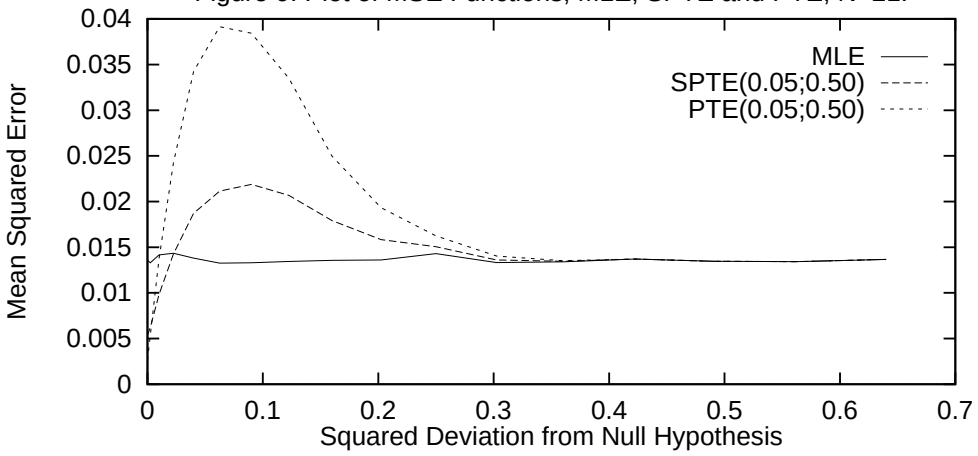
α Level	Sample N_1 Value	Sample N_2 Value	Sample N_3 Value	Sample N_4 Value	Asymptotic Value
0.15	7.0488	5.1888	4.5878	4.2945	2.0723
0.10	10.154	7.0539	6.0830	5.6468	2.7055
0.05	17.2416	10.8074	8.9003	8.1403	3.8415
0.01	48.0163	22.9328	16.5186	14.7270	6.6349
0.005	71.6056	29.6674	20.3324	17.9997	7.8795
0.001	173.6875	50.7775	30.8321	26.4770	10.8280

Figure 8: Plot of MSE Functions, MLE, SPTE and PTE, N=11.



The horizontal axis has as its scale δ^2, so is not normalised by σ^2. As expected, the pattern of the plots of $\hat{\mu}_p^{SP}$ and $\hat{\mu}_p^{P}$ is what one sees in the analytical results plotted in Figure 4 earlier. That is, $\hat{\mu}_p^{SP}$ quickly dominates $\hat{\mu}_p^{P}$ in terms of *MSE* for small departures from $\delta^2 = 0$, and cuts $MSE(\hat{\mu}_p)$ at around $\delta = 0.25$. This is large relative to the precision of $\hat{\mu}_p$, which has a standard error of 0.01 across all 500,000 sample values. Also, for $\mu_{p_o} \pm 0.25$, the associated "kill rates" would be 0.622 and 0.378 respectively. Thus, if a researcher intends to identify ED_{50} with minimum risk, then $\hat{\mu}_p^{SP}$ provides the most appropriate estimator to use when D indicates deviations from H_0 in this neighbourhood of δ.

Figure 9: Plot of MSE Functions, MLE, SPTE and PTE, N=21.



One means by which to improve the performance of $\hat{\mu}_p^{SP}$ and $\hat{\mu}_p^{P}$ is for a researcher to make use of a larger sample. Figure 9 provides a comparison of $\hat{\mu}_p^{SP}$ and $\hat{\mu}_p^{P}$ to $\hat{\mu}_p$ at $N_2 = 21$, again with $\pi = 0.50$ and at $\alpha = 0.05$. The superior performance of $\hat{\mu}_p^{SP}$ relative to $\hat{\mu}_p^{P}$ is again evident for small deviations from $\delta = 0$. However, it is now relatively more easy to discriminate quickly with regard to when a departure from H_o has occurred. That is, $MSE(\hat{\mu}_p^{SP})$ cuts $MSE(\hat{\mu}_p)$ at $\delta = 0.15$. This translates into associated kill rates for $\mu_{p_o} \pm 0.15$, of 0.575 and 0.425 respectively. Thus, the degree of imprecision or risk associated with estimating ED_{50} is substantially reduced.

As could be expected, the precision in estimation of ED_{50} can be further improved by increasing sample sizes still further. Figures 10 and 11 indicate the *MSE* functions for $\hat{\mu}_p^{SP}, \hat{\mu}_p^{P}$ and $\hat{\mu}_p$ at $N_3 = 50$ and $N_4 = 105$ respectively. However, while increasing the sample size from $N_2 = 21$ to $N_3 = 50$ causes a reduction in the point where $MSE(\hat{\mu}_p^{SP})$ cuts $MSE(\hat{\mu}_p)$ from $\delta = 0.15$ to $\delta = 0.08$, the further increase of $N_3 = 50$ to $N_4 = 105$ only reduces this intersection point to $\delta = 0.07$. Thus, from a practical point of view for an applied researcher, the benefit of increasing sample size beyond $N_3 = 50$ in terms of estimator precision will more than likely be out-weighed by the added cost of additional sampling. Depending on the nature of the application, this increased cost is likely to be non-trivial, whether in pecuniary, environmental or human terms.

Figure 10: Plot of MSE Functions, MLE, SPTE and PTE, N=50.

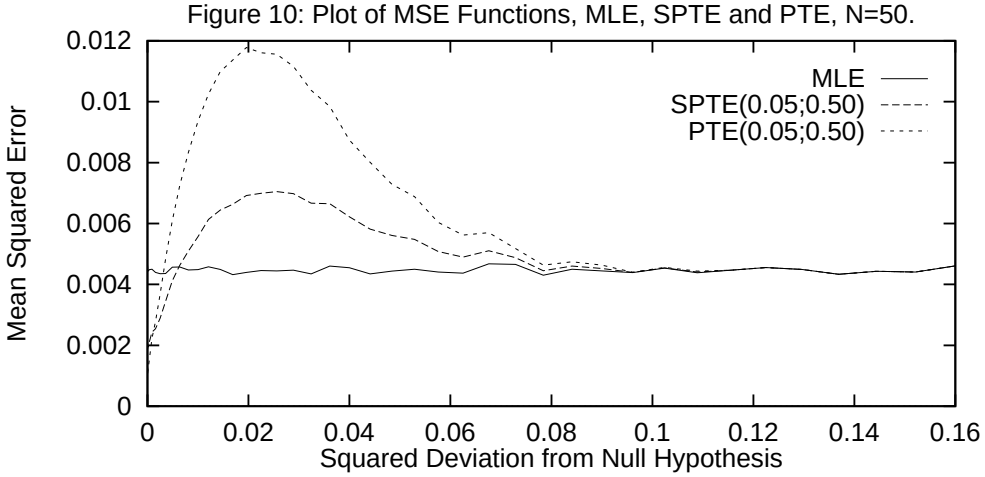
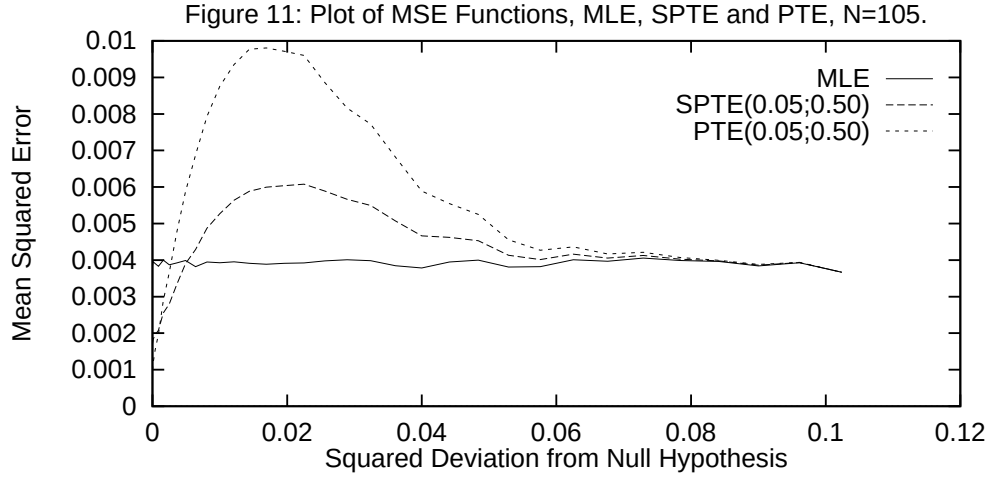


Figure 11: Plot of MSE Functions, MLE, SPTE and PTE, N=105.



In Section 5, the notion of *ARE* was defined, and used to illustrate how an applied researcher could select optimal values of α and π. We conclude this Section by presenting results pertaining to the *simulated relative efficiency* (*SRE*) of $\hat{\mu}_p^{SP}$, based on the simulation experiment with $N_3 = 50$ observations. Table 4 contains these results, where the *SRE* are computed according to the definition of *ARE* specified in (5.1), which was based on *AMSE*, rather than *simulated mean squared error* (*SMSE*), on which *SRE* is based.

This Table includes *SRE* for a series of values of δ, up to the point roughly where $SMSE(\hat{\mu}_p^{SP})$ cuts $SMSE(\hat{\mu}_p)$. Values of *SRE* are reported for three values of $\pi, 0.25, 0.50$ and 0.75; and for three values of $\alpha, 0.10, 0.05$ and 0.01. At values of $\delta = 0.08$, there is still substantial improvement in the performance of $\hat{\mu}_p^{SP}$ over $\hat{\mu}_p$ in terms of *SMSE*, the criterion on which *SRE* is implicitly based. This improvement is the case as an increasing function of π, and as a decreasing function of π. In addition, the sample size at $N_3 = 50$ is on the boundary of what one could reasonably expect to use in applied situations. For smaller sample sizes, the *MSE* gains in using $\hat{\mu}_p^{SP}$ over $\hat{\mu}_p$ would accrue over an even greater range of values for δ. Thus, $\hat{\mu}_p^{SP}$ provides an extremely useful way in which *UPI* can be used in applied settings such as this.

7 Summary and Concluding Remarks

In this paper, three shrinkage estimators are presented, as alternatives to the usual maximum likelihood estimator *MLE* for estimation of mortality rates for a given dosage. Probabilities of mortality are assumed to follow a logistic distribution function, which is a commonly used model in this literature.

The focus is to obtain alternative estimators of ED_{50}, the dosage associated with a 50% mortality rate. The three estimators presented as alternatives to the *MLE*, $\hat{\mu}_p$, are a shrinkage estimator based on incorporation of uncertain prior information (*UPI*), denoted $\hat{\mu}_p^S$, a typical pre-test estimator, $\hat{\mu}_p^P$, and a shrinkage preliminary test estimator, $\hat{\mu}_p^{SP}$. We show that the $\hat{\mu}_p^{SP}$ displays superior performance to the other three estimators for a significant range of the parameter space. On the basis of simulation experiment results, this range appears to be substantial, in terms of yielding large reductions in mean squared error. In fact, the

Table 4: Comparisons of Simulated Relative Efficiencies (SRE)
of $\hat{\mu}_p^{SP}$ at $\pi = 0.25$, 0.50 and 0.75,
Various Values of α.

	SRE at $\alpha = 0.10$			SRE at $\alpha = 0.05$			SRE at $\alpha = 0.01$		
δ	$\pi = 0.25$	$\pi = 0.50$	$\pi = 0.75$	$\pi = 0.25$	$\pi = 0.50$	$\pi = 0.75$	$\pi = 0.25$	$\pi = 0.50$	$\pi = 0.75$
0.00	1.369	1.858	2.366	1.515	2.397	3.683	1.713	3.488	9.227
0.01	1.369	1.848	2.317	1.523	2.410	3.642	1.715	3.464	8.550
0.02	1.347	1.763	2.111	1.506	2.304	3.226	1.739	3.537	8.138
0.03	1.312	1.643	1.852	1.475	2.135	2.692	1.710	3.203	5.498
0.04	1.298	1.576	1.687	1.477	2.074	2.411	1.679	2.881	3.920
0.05	1.231	1.388	1.383	1.426	1.844	1.911	1.666	2.646	3.000
0.06	1.180	1.259	1.191	1.345	1.578	1.491	1.618	2.304	2.230
0.07	1.120	1.135	1.037	1.277	1.391	1.241	1.584	2.072	1.807
0.08	1.076	1.045	0.923	1.200	1.214	1.028	1.549	1.848	1.456
0.09	1.026	0.951	0.812	1.141	1.081	0.874	1.473	1.569	1.142
0.10	0.988	0.892	0.754	1.074	0.969	0.768	1.405	1.374	0.954
0.11	0.950	0.837	0.702	1.012	0.874	0.682	1.326	1.196	0.803
0.12	0.922	0.793	0.653	0.970	0.806	0.614	1.243	1.029	0.666
0.13	0.895	0.757	0.621	0.919	0.738	0.555	1.168	0.902	0.569
0.14	0.881	0.743	0.612	0.884	0.701	0.532	1.088	0.801	0.503
0.15	0.880	0.748	0.623	0.864	0.681	0.517	1.054	0.742	0.456

smaller the sample sizes used by an applied researcher, the greater the potential gains in using $\hat{\mu}_p^{SP}$. The results of our simulation experiments also indicate several factors which an applied researcher would wish to take into account in practical situations. In particular, the results have implications for choice of sample sizes, caution in the use of the asymptotically valid test statistic used in the construction of $\hat{\mu}_p^{SP}$ and $\hat{\mu}_p^{P}$, and the extent to which one can gain in the use of, say, $\hat{\mu}_p^{SP}$, for deviations of μ_{p_o} from μ_p.

8 References

Ahmed, S. E. (1997), "Asymptotic Shrinkage Estimation: The Regression Case." In *Applied Statistical Science, II*, Papers in Honor of Munir Ahmad, M. Ahsanullah, Editor. Nova Science Publishers Inc.

Bancroft, T. A. (1944), "On biases in estimation due to the use of preliminary tests of significance." *Annals of Mathematical Statistics*, **15**, 190–204.

Goedhart, P. W. (1986), "Statistical Analysis of Quantal Assay Data." ITI – TNO Report No. 86, ITI A 31.

Hamilton, M. A. (1979), " Robust Estimates of the ED50". *Journal of the American Statistical Association*, **74**, 344–354.

Hoekstra, J. A. (1991), "Estimation of the LC50, A Review." *Environmetrics*, **2**, 139–152.

Judge, G. G. and M. E. Bock (1978), *The Statistical Implications of Pre-test and Stein-rule Estimators in Econometrics*. North Holland, Amsterdam.

White, K. J. (1978), "A General Computer Program for Econometric Methods – SHAZAM". *Econometrica*, **46**, 239–240.

Zellner, A. and T. H. Lee (1965), "Joint Estimation of Relationships Involving Discrete Random Variables". *Econometrica*, **33**, 382–394.

Appendix

Table A1: Comparisons of Maximum and Minimum $ARE(\hat{\mu}_p^{SP})$, $\pi = 0.0(0.1)1.0$, and for Various Values of α.

						$\alpha = 0.01$					
π	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
Δ_o	0.0	8.0	7.0	7.0	7.0	7.0	7.0	7.0	6.0	6.0	6.0
E_o	1.000	0.758	0.590	0.469	0.380	0.313	0.262	0.222	0.190	0.164	0.143
E^*	1.000	1.211	1.492	1.876	2.415	3.191	4.330	5.992	8.257	10.678	11.836
						$\alpha = 0.05$					
π	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
Δ_o	0.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	4.0	4.0
E_o	1.000	0.833	0.701	0.596	0.512	0.443	0.387	0.341	0.302	0.269	0.240
E^*	1.000	1.159	1.350	1.581	1.857	2.177	2.535	2.907	3.247	3.493	3.583
						$\alpha = 0.10$					
π	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
Δ_o	0.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
E_o	1.000	0.874	0.768	0.678	0.602	0.537	0.481	0.434	0.392	0.356	0.325
E^*	1.000	1.119	1.253	1.400	1.560	1.726	1.890	2.042	2.166	2.248	2.276
						$\alpha = 0.15$					
π	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
Δ_o	0.0	4.0	4.0	4.0	4.0	4.0	4.0	3.0	3.0	3.0	3.0
E_o	1.000	0.901	0.815	0.739	0.672	0.613	0.561	0.514	0.472	0.434	0.401
E^*	1.000	1.092	1.189	1.291	1.395	1.497	1.592	1.674	1.738	1.779	1.794

Table A1: Comparisons of Maximum and Minimum $ARE(\hat{\mu}_p^{SP})$, $\pi = 0.0(0.1)1.0$, and for Various Values of α, continued.

						$\alpha = 0.20$					
π	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
Δ_o	0.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
E_o	1.000	0.923	0.852	0.787	0.728	0.675	0.626	0.582	0.542	0.505	0.472
E^*	1.000	1.071	1.144	1.217	1.289	1.356	1.417	1.468	1.506	1.531	1.539
						$\alpha = 0.25$					
π	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
Δ_o	0.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
E_o	1.000	0.938	0.880	0.826	0.776	0.729	0.686	0.646	0.608	0.574	0.542
E^*	1.000	1.055	1.110	1.164	1.215	1.262	1.302	1.336	1.361	1.377	1.382
						$\alpha = 0.30$					
π	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
Δ_o	0.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
E_o	1.000	0.951	0.904	0.860	0.818	0.778	0.741	0.705	0.672	0.640	0.610
E^*	1.000	1.043	1.085	1.124	1.161	1.194	1.223	1.246	1.263	1.273	1.277
						$\alpha = 0.35$					
π	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
Δ_o	0.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0
E_o	1.000	0.962	0.925	0.889	0.855	0.822	0.790	0.759	0.730	0.702	0.675
E^*	1.000	1.033	1.064	1.094	1.121	1.144	1.165	1.181	1.193	1.200	1.202