CentER for Economic Research

Discussion paper

[redacted]

TILBURG UNIVERSITY

No. 9165

SELECTING ESTIMATED MODELS USING CHI-SQUARE STATISTICS

by Quang H. Vuong
and Weiren Wang

December 1991

ISSN 0924-7815

SELECTING ESTIMATED MODELS USING CHI-SQUARE STATISTICS*

Quang H. Vuong
INRA and University of Southern California

and

Weiren Wang
University of Kentucky

October 1991

* Correspondence: Professor W. Wang
Dept. of Economics
College of Business and Economics
University of Kentucky
Lexington, KY 40506-0034
U.S.A.

SELECTING ESTIMATED MODELS USING CHI-SQUARE STATISTICS

Quang H. Vuong
INRA and University of Southern California
and
Weiren Wang
University of Kentucky

ABSTRACT

This paper proposes some tests for choosing estimated models from two competing parametric families using Pearson type statistics. We allow arbitrary asymptotically normal estimators to be used in forming Pearson type goodness-of-fit statistics. In practice such a construction provides flexibility in applying our tests. Large Sample theory and bootstrap methods are used to construct our tests.

Keywords: Model Selection, Chi Square Statistics, Bootstrap, Misspecified Models, Nonnested Hypotheses.

Acknowledgement: We would like to thank C. Hsiao, A. Weiss, two referees, seminar participants at the Unniversity of Kentucky, Free University and Tilburg University for helpful comments. The second author wants to express his gratitude to a fellowship from CentER for Economic Research at Tilburg University for partial financial support.

SELECTING ESTIMATES USING CHI-SQUARE STATISTICS

Quang H. Vuong and Weiren Wang

1. INTRODUCTION

Pearson (1900) type chi-square statistics have been generally used to test goodness-of-fit, i.e., to test whether a specified parametric model is consistent with the observed data. Cochran (1952), Watson (1959), and Moore (1978,1986) have provided comprehensive surveys on Pearson chi-square type statistics, i.e., quadratic forms in the cell frequencies. Recently, Andrews (1988a,1988b) has extended the Pearson chi-square testing method to non-dynamic parametric econometric models, i.e., to models with covariates. Because Pearson chi-square statistics provide natural measures for the discrepancy between the observed data and a specific parametric model, they have also been used for discriminating among competing models. Such a situation is frequent in Social Sciences where many competing models are proposed to fit a given sample. A well known difficulty is that each chi-square statistic tends to become large without an increase in its degrees of freedom as the sample size increases. As a consequence goodness- of-fit tests based on Pearson type chi-square statistics will generally reject the correct specification of every competing model.

To circumvent such a difficulty, a popular method for model selection, which is similar to the use of the Akaike (1973) Information Criterion (AIC), consists in considering that the lower the chi-square statistic, the better is the model. Hence the parametric model with smaller value of chi-square statistic will be chosen. Such a use of chi-square statistics has been suggested by various researchers. See Massy, Montgomery and Morrison (1970), Heckman (1981), and Nakamura and Nakamura (1985) among others.

The preceding selection rule, however, is not entirely satisfactory. Since chi-square statistics depend on the sample and are therefore random, their actual values are subject to statistical variations. Hence a model with a smaller chi-square statistic is not necessarily better than one with a larger chi-square statistic in terms of goodness-of-fit. To take into account statistical variations, we shall propose some convenient asymptotically standard normal tests for model selection based on Pearson type chi-square statistics. Following Vuong (1989) our tests are testing the null hypothesis that the competing models are as close to the data generating process (DGP) against the alternative hypotheses that one model is closer to the DGP where closeness of a model is measured according to the discrepancy implicit in the Pearson type chi-square statistics. Thus the outcomes of our tests provide information on the strength of the statistical evidence for the choice of a model based on its goodness-of-fit.

Following Moore (1977,1978) and Andrews (1988b), we consider a general class of estimators that is very broad and contains most estimators currently used in practice, when forming Pearson type statistics. This covers the case studied in Vuong and Wang (1990) where only the corresponding minimum chi-square estimator is used. In practice, the use of "unmatched" estimators, i.e., estimators that do not minimimize the chosen Pearson type statistic, is quite common and can be found in various works. See Linhart and Zucchini (1986), Heckman (1981), and Nakamura and Nakamura (1985). Such a generalization is useful for many reasons.

First, there frequently exist estimators that are easier to compute than the corresponding minimum chi-square estimator. Second, when the original individual data are available, one is tempted to use more efficient estimators, such as the maximum likelihood estimator (MLE) based on the individual data when forming the chi-square type statistic. This is in fact the procedure recommended in many textbooks. See for instance, Bishop, Fienberg and Holland (1975). Third, a particular chi-square statistic is often chosen to obtain a chi-square limiting null distribution for a given estimation method. For example, the Rao-Robson statistic has the χ^2_{M-1} limiting null distribution when the raw MLE is used. However, the raw MLE does not minimize the Rao-Robson statistic. See Chernoff and Lehmann (1954), Moore and Spruill (1975), Moore (1977, 1978) and Andrews (1988a).

The paper is organized as follows. Section 2 introduces the basic notations and defines the class of asymptotically normal (AN) estimators used. Section 3 investigates the model selection problem based on our general Pearson type statistics. In section 4, methods based on the boostrap are used to propose alternative testing procedures for model selection. Section 5 presents some simulation results. Section 6 concludes the paper and briefly mentions directions for future research. Some proofs and tables are included in Appendix.

2. DEFINITIONS AND ASSUMPTIONS

In this section, we briefly present the basic assumptions on the model and parameter estimators, and we define our general chi-square type statistics.

Assumption A1: The observed data X_i, $i=1,2,\ldots$ are independent and identically distributed (iid) with some common true distribution H.

The sample space X is partitioned into M mutually disjoint fixed cells $E_1, E_2, \ldots, E_M$. The partition is sometimes based either on the nature of the data, say race and sex, or survey designer. Let $h = (h_1, h_2, \ldots, h_M)$ be the vector of true cell probabilities. Let a specified model be $F_\theta = \{F(.|\theta); \theta\in\Theta\}$ and denote the vector of its predicted cell probabilities by $p(\theta)=[p_1(\theta), p_2(\theta),\ldots,p_M(\theta)]'$ where $p_i(\theta)= \int_{Ei} dF(x|\theta)$. Throughout the paper, we impose the following assumption on h and $p(\theta)$:

Assumption A2: $h_i>0$, $p_i(\theta)>0$, and $p_i(\theta)$ is twice continuously differentiable for every $i=1,2,\ldots,M$.

Throughout it is assumed that F_θ satisfies some standard additional regularity conditions to ensure the asymptotic results presented subsequently. See for instance Moore and Spruill (1975) and Moore (1984).

Let n be the sample size. Corresponding to the partition $E_1, E_2,\ldots, E_M$, we can compute the vector of observed cell probabilities

$$f = (f_1,f_2,\ldots,f_M)' \text{ where } f_i= \frac{1}{n} \sum_{j=1}^{n} 1_{E_i}(X_j) \quad \text{for } i=1,2,\ldots,M,$$

and $1_{E_i}(X_j)$ is the indicator function taking values 1 if X_j falls in cell E_i or 0 otherwise. Following Moore (1978), it is convenient to consider the M-dimensional vector

$$V_n(\theta) = \surd n(\ldots,\frac{f_i-p_i(\theta)}{\surd p_i(\theta)},\ldots)' ,$$

which measures the standardized difference between the observed and the expected cell probabilities given θ.

We are in a position to define the class of general chi-square type statistics considered in this paper. These statistics are essentially quadratic forms in the standardized cell frequencies $V_n(\theta_n)$. Formally, we have:

Definition 1: A general chi-square type statistic is of the form

$$Q_n(\theta_n) = V_n'(\theta_n)M(f,\theta_n)V_n(\theta_n)$$

where $M(f,\theta)$ and θ_n satisfy the following assumptions A3 and A4, respectively.

Assumption A3: Each element of the weighting matrix $M(f,\theta)$ is twice continuously differentiable in $(f,\theta) \in R^M \times \Theta$, and $M(h,\theta)$ is a positive definite matrix for every θ.

In this paper, we consider estimators that satisfy the next regularity assumption.

Assumption A4: For some θ_o in Θ, the estimator θ_n satisfies

$$\sqrt{n}(\theta_n-\theta_o) = \frac{1}{\sqrt{n}} R_o^{-1} \sum_{i=1}^{n} \psi(X_i,\theta_o) + o_p(1) \quad \text{as } n \dashrightarrow \infty$$

where $\psi(X,\theta_o)$ is a measurable function from $\mathbf{X} \times \Theta$ to R^k that satisfies $E_H\psi(X,\theta_o)=0$, and $V_o = E_H\psi(X,\theta_o)\psi(X,\theta_o)'$ and $R_o = -E_H(\partial\psi(X,\theta_o)/\partial\theta')$ are finite and nonsingular. $E_H(.)$ denotes expectation computed under the true data generating process.

Assumption A4 implies that θ_n is a consistent estimator of some value θ_o and that $\sqrt{n}(\theta_n-\theta_o)$ is asymptotically normally distributed with zero mean and covariance matrix $R_o^{-1}V_o(R_o^{-1})'$. Given some suitable regularity conditions, most common estimators θ_n fulfill this assumption. For instance, the minimum chi-square estimators, the maximum likelihood (ML) estimator on grouped or ungrouped data, any GMM estimator, and other extremum estimators satisfy assumption A4. See Amemiya (1985).

Note that the parameter value θ_o depends on the underlying true distribution H which generates the observations as well as the estimation method employed. For example, when the minimum chi-square estimator is used, then θ_o is actually minimizing $Q(\theta) = \text{plim } Q_n(\theta)/n$. On the other hand, if θ_n is the MLE of θ based on the ungrouped sample data, then θ_o is the value at which the Kullback-Leibler information criterion (KLIC) $E_H[-\log f(x|\theta)]$ is minimized, where f is the density function corresponding to $F(.|\theta)$. See White (1982). In general, when the model F_θ is misspecified, these two values for θ_o are different.

Note that $Q_n(\theta_n)$ is very general. It includes some well-known chi-square statistics such as the Pearson statistic with $M_n = I_M$ (the identity matrix), the Modified Pearson statistic with $M_n = \text{diag}(\ldots,p_i(\theta)/f_i,\ldots)$, the Gauss statistic with $M_n = \text{diag}(\ldots,p_i(\theta),\ldots)$ and the Rao-Robson statistic with M_n being the generalized inverse of the covariance matrix of $V_n(\theta_n)$. In practice, M_n is specified by researcher according to his or her preference or objective.

3. SELECTING ESTIMATESD MODELS

As we mentioned earlier, chi-square statistics are frequently used to discriminate among alternative models. It is easy to see that, under the present regularity conditions, $Q_n(\theta)/n$ converges to $Q(\theta)=V(\theta)'M(h,\theta)V(\theta)$ in probability as n goes to infinity, where $V(\theta)=[\ldots,(h_i-p_i(\theta))/\sqrt{p_i(\theta)},\ldots]'$. Thus $Q(\theta)$ can be viewed as measuring the departure of a particular member $F(.|\theta) \in F_\theta$ from the observed sample. It is also worth noting that $Q(\theta)\geq 0$ and $Q(\theta)=0$ if and only if $h=p(\theta)$.

Of special interests to us is the situation in which a researcher has two competing parametric models F_θ and G_γ, and desires to select the better of the two models based on their general chi-square type statistics $Q_n(\theta_n)$ and $Q_n(\gamma_n)$ where θ_n and γ_n are general estimators satisfying assumption A4 and the same cells are used in both statistics.

Definition 2 (Equivalent, Better and Worse): Given two competing parametric models F_θ and G_γ and some chi-square type statistics $Q_n(\theta_n)$ and $Q_n(\gamma_n)$ where θ_n and γ_n are general estimators satisfying A4,

$\tilde{H}_o^e$: $Q(\theta_o) = Q(\gamma_o)$ means that the two models are equivalent,

$\tilde{H}_f$: $Q(\theta_o) < Q(\gamma_o)$ means that model F_{θ_o} is better than G_{γ_o},

$\tilde{H}_g$: $Q(\theta_o) > Q(\gamma_o)$ means that model F_{θ_o} is worse than G_{γ_o}.

where $Q(.)$ is the probability limit of $Q_n(.)/n$.

Definition 2 allows the use of estimators other than the corresponding minimum chi-square estimators. It does not even require that the same chi-square type discrepancy be used in forming $Q_n(\theta_n)$ and $Q_n(\gamma_n)$. Note, however, that choosing different chi-square discrepancies for evaluating competing models is hardly justified.

More importantly, even if the same chi-square discrepancy $Q(.)$ is used, it is important to note that the preceding hypotheses are not entirely consistent with the problem of model selection which is that of choosing between the models F_θ and G_γ and not between the estimates $F(.|\theta_n)$ and $G(.|\gamma_n)$. This is so because the probability limits θ_o and γ_o of the estimators θ_n and γ_n are not in general equal to the pseudo true values θ_* and γ_* associated with the discrepancy $Q(.)$ ($\theta_* = \text{argmin } Q(\theta)$ and $\gamma_* = \text{argmin } Q(\gamma)$). As a consequence $Q(\theta_o)$ (say) cannot be interpreted as the discrepancy between the model F_θ and the DGP H. Thus the preceding hypotheses are not only properties of the competing models but also of the estimation methods employed. The relative effects due to specification errors and choice of estimation are not exactly known. However, if θ_n and γ_n are minimum chi-square estimators, then one can analyze the effect due to specification errors using procedures developed in Vuong and Wang (1990).

Nonetheless, consideration of the general hypotheses $\tilde{H}_o^e$, $\tilde{H}_f$, and $\tilde{H}_g$ is useful for two reasons. First, "unmatched" estimators are quite common in practice for reasons mentioned in the introduction section. Secondly, one may be interested only in comparing the performance of the estimated models $F(.|\theta_n)$ and $G(.|\gamma_n)$. Then $\tilde{H}_o^e$, $\tilde{H}_f$, and $\tilde{H}_g$ are the relevant hypotheses.

In any case, since θ_n and γ_n are consistent estimators of θ_o and γ_o by assumption A4, we can use $[Q_n(\theta_n)-Q_n(\gamma_n)]/n$ to consistently estimate the indicator $Q(\theta_o)-Q(\gamma_o)$ which will be zero under the null hypothesis $\tilde{H}_o^e$. Following standard exercise, we obtain the asymptotic distribution of $[Q_n(\theta_n)-Q_n(\gamma_n)]/\sqrt{n}$, which is normal distribution with zero mean and variance ω^2. The formula for ω^2 and detailed derivation can be found in Appendix. Hence we define the statistic

$$GCM_n = \frac{1}{\sqrt{n}} \frac{[Q_n(\theta_n) - Q_n(\gamma_n)]}{\hat{\omega}} .$$

where $\hat{\omega}^2$ is a consistent estimator of ω^2 (e.g., its sample analog). We have

Theorem 3.2 (Asymptotic Distribution of GCM_n Statistic): Given A1-A4, suppose that $\omega^2 \neq 0$, then

(i) under the null hypothesis $\tilde{H}_o^e$, $GCM_n \to N(0,1)$ in distribution,

(ii) under the alternative $\tilde{H}_f$, $GCM_n \to -\infty$ in probability,

(iii) under the alternative $\tilde{H}_g$, $GCM_n \to +\infty$ in probability.

Theorem 3.2 is quite general and gives us a wide variety of asymptotic standard normal tests for model selection based on general chi-square type statistics. Part (ii) and (iii) also implies that the test is consistent. In the next section, we detail the testing procedures based on Theorem 3.2 by using bootstrap methods.

The above tests are based on the standardized difference in *unadjusted* generalized chi-square statistics. In some cases, especially when the sample size is small, one may want to add adjusting factors like $(p-q)/\sqrt{n}$ to our test statistics. Similar results are expected to hold. However, as the sample size becomes very large, the effect of the adjusting factors will vanish.

4. BOOTSTRAP METHODS

Implementation of the model selection procedure proposed in section 3 requires the following computations:

(i) Estimation of the parameters θ_n and γ_n,

(ii) Computation of the two chi-squares statistics $Q_n(\theta_n)$ and $Q_n(\gamma_n)$ and the difference $\hat{S}_n = [Q_n(\theta_n)-Q_n(\gamma_n)]/\sqrt{n}$,

(iii) Computation of the variance $\hat{\omega}^2$ of $\hat{S}_n$ and finally, computation of $GCM_n = \hat{S}_n/\hat{\omega}$.

The estimators θ_n and γ_n can be obtained by minimizing some objective function, such as $Q_n(.)$, or by maximizing the likelihood function. Point (ii) is straightforward once θ_n and γ_n are known. Point (iii) is somewhat complicated. In particular, the asymptotic formula for calculating the variance $\hat{\omega}^2$ involves the calculation of the first and second order partial derivatives of the expected cell probabilities with respect to the parameters θ and γ. Moreover, such a formula varies across models and estimation methods. This is not very convenient in applied work. Fortunately, with the help of advanced computers, $\hat{\omega}^2$ can be obtained via some simulation techniques. Specifically, we shall consider a method for evaluating ω^2 based on the bootstrap method. In addition, we shall propose two alternative testing procedures for model selection based directly on the bootstrap distribution of the statistic $\hat{S}_n$.

In the preceding section, we have seen that $\hat{S}_n$ is approximately normally distributed with mean zero and variance ω^2. This suggests that ω^2 can be estimated by the sample variance of $\hat{S}_n$ in a (large) number of independent and identical samples of size n. This is the basic idea underlying the bootstrap method which we apply here to the estimation of the variance $\hat{\omega}^2$. Specifically, we carry out the following steps:

1. For each set of data $x_1, x_2, \ldots, x_n$, let $\hat{F}$ be the empirical probability distribution of the data, i.e.,

$$\hat{F}\text{: mass } 1/n \text{ at } x_i, \quad i=1,2,\ldots,n.$$

Then draw an i.i.d. "bootstrap sample" $x_1^*, x_2^*, \ldots, x_n^*$ from $\hat{F}$, i.e., draw x_i^* randomly with replacement from the observed values $x_1, x_2, \ldots, x_n$.

2. Using this bootstrap sample $\{x_i^*\}$, estimate the competing models to obtain θ_n^* and γ_n^*. Then calculate the statistic

$$\hat{S}_n^* = [Q_n(\theta_n^*) - Q_n(\beta_n^*)]/\sqrt{n} \ .$$

3. Independently repeat steps 1 and 2 a large number of times B, say B=1000. Obtain "bootstrap replications" $\hat{S}_n^{*1}, \hat{S}_n^{*2}, \ldots, \hat{S}_n^{*B}$, and compute the sample variance of $\{\hat{s}_n^{*b}, b=1,\ldots,B\}$:

$$\hat{\omega}_*^2 = \frac{1}{B} \sum_{b=1}^{B} (\hat{S}_n^{*b} - \bar{S}^*)^2 ,$$

where $\bar{S}^* = \frac{1}{B} \sum_{b=1}^{B} \hat{S}_n^{*b}$ is the average of "bootstrap replications".

The above method (call it Method 1) is quite general. If only the frequencies in every cells are available instead of the individual data so that estimation methods based on grouped data must be used, the bootstrap method can still be applied with a slight modification of the resampling procedure. Specifically, the following modified procedure (Method 1') can be used:

1'. Given the observed cell probabilities $(f_i, i=1,\ldots,M)$ and the total sample size n, we can construct artificial data $a_1, a_2,\ldots,a_n$, such that it has the same observed cell probabilities. For example, we can draw nf_i points from uniform distribution in cell i, and treat these nf_i points as if they were true sample observations in cell i.

2' Now we treat the above artificial sample $a_1, a_2,\ldots,a_n$ as if they were original true sample $x_1, x_2,\ldots,x_n$ and repeat three steps in Method 1.

Once the bootstrap variance $\hat{\omega}_*^2$ is obtained, the test statistic GCM_n is calculated easily using the initial estimates θ_n and γ_n. Under suitable regularity conditions and for a large number of replications (see Efron (1982)), $\hat{\omega}_*^2$ is a consistent estimator of ω^2. Thus, from Theorem 3.2, a testing procedure for model selection can be based on the comparison of the value of GCM_n to critical values from a standard normal table. For example, at 5% significance level, we compare GCM_n with -1.96 and 1.96. If GCM_n falls between -1.96 and 1.96, we conclude that both estimated models fit the data equally well. If GCM_n is less than -1.96 (or larger than 1.96), then we reject the null hypothesis in favor of the alternative hypothesis that the estimated model $F(.|\theta_n)$ (or $G(.|\gamma_n)$) is closer to the true distribution.

Although using the bootstrap method to obtain an estimate of ω^2, the basic justification of the preceding testing procedure comes from the asymptotic properties obtained in Theorem 3.2. In contrast, the next testing procedures rely only on the bootstrap methodology, and in particular on two bootstrap methods for assigning approximate confidence intervals to $Q(\theta_o) - Q(\gamma_o)$ based on the bootstrap distribution of $\hat{S}_n^*$. These two methods have been discussed in detail in Efron (1982), and require steps 1 and 2 discussed above or steps 1′ and 2′ if only frequencies are observed. See also Efron (1984) for the comparison of nonnested linear models using a MSE criterion.

The first testing method is based on the percentile method. Let

$$\hat{CDF}(t) = \frac{\text{number of } \{b\colon \hat{S}_n^{*b} \le t\}}{B} = \text{Prob}_*(\hat{S}_n^* \le t)$$

be the empirical cumulative distribution function (CDF) of the bootstrap distribution of $\{\hat{S}_n^{*b},\ b=1,\ldots,B\}$. For a given significance level α between 0.0 and 1.0, define

$$\hat{S}_L^*(\alpha/2) = \hat{CDF}^{-1}(\alpha/2) \quad \text{and} \quad \hat{S}_U^*(\alpha/2) = \hat{CDF}^{-1}(1-\alpha/2) \ .$$

The percentile method consists in taking $[\hat{S}_L^*(\alpha/2), \hat{S}_U^*(\alpha/2)]$ as an approximate $1-\alpha$ central confidence interval for $Q(\theta_o) - Q(\gamma_o)$. Thus a test of the null hypohtesis $\tilde{H}_o^e$ of equivalence against the alternative hypothesis $\tilde{H}_f$ or $\tilde{H}_g$ at the approximate α significance level is:

(i) accept the null $\tilde{H}_o^e$ of equivalence if $0 \in [\hat{S}_L^*(\alpha/2), \hat{S}_U^*(\alpha/2)]$,

(ii) reject $\tilde{H}_o^e$ in favor of $\tilde{H}_f$ if $\hat{S}_U^*(\alpha/2) < 0$,

(iii) reject $\tilde{H}_o^e$ in favor of $\tilde{H}_g$ if $\hat{S}_L^*(\alpha/2) > 0$.

Or equivalently,

(i) accept $\tilde{H}_o^e$ if $\hat{CDF}(0) \in [\alpha/2,\ 1-\alpha/2]$,

(ii) reject $\tilde{H}_o^e$ in favor of $\tilde{H}_f$ if $\hat{CDF}(0) > 1-\alpha/2$,

(iii) reject $\tilde{H}_o^e$ in favor of $\tilde{H}_g$ if $\hat{CDF}(0) < \alpha/2$.

The percentile method does not use the value $\hat{S}_n$ for the initial observed sample. More importantly, since the bootstrap distribution is based on replications of the observed sample that produces the value $\hat{S}_n$, the precentile method assumes implicitly that $\hat{S}_n$ is the median of the bootstrap distribution. If this is not a proper assumption, one should incorporate a bias adjustment. This leads to the bias-correction percentile method. We will only present the procedure. Its rationale can be found in Chapter 10 of Efron (1982).

Define

$$\hat{z} = \Phi^{-1}(\widehat{CDF}(\hat{S}_n)) \text{ and } z_{\alpha/2} = \Phi^{-1}(1-\alpha/2)$$

where Φ is the cumulative distribution function for a standard normal variable. The decision rule for model selection based on the bias-correction percentile method at the α significance level is:

(i) accept $\tilde{H}_o^e$ if $0 \in [\widehat{CDF}^{-1}(\Phi(2\hat{z} - z_{\alpha/2})), \widehat{CDF}^{-1}(\Phi(2\hat{z} + z_{\alpha/2}))]$,

(ii) reject $\tilde{H}_o^e$ in favor of $\tilde{H}_f$ if $\widehat{CDF}^{-1}(\Phi(2\hat{z} + z_{\alpha/2})) <0$,

(iii) reject $\tilde{H}_o^e$ in favor of $\tilde{H}_g$ if $\widehat{CDF}^{-1}(\Phi(2\hat{z} - z_{\alpha/2})) >0$.

Let $\hat{z}_0 = \Phi^{-1}(\widehat{CDF}(0))$. It is easy to see that the preceding decision rule is equivalent to:

(i) accept $\tilde{H}_o^e$ if $|\hat{z}_0 - 2\hat{z}| < z_{\alpha/2}$,

(ii) reject $\tilde{H}_o^e$ in favor of $\tilde{H}_f$ if $\hat{z}_0 > 2\hat{z} + z_{\alpha/2}$,

(iii) reject $\tilde{H}_o^e$ in favor of $\tilde{H}_g$ if $\hat{z}_0 < 2\hat{z} - z_{\alpha/2}$.

We will refer to the percentile method without the bias-correction as Method 2, and to the percentile method with the bias-correction as Method 3.

5. AN EXAMPLE

To illustrate the model selection procedures discussed in the preceding section, namely, Methods 1, 2 and 3, we consider an example. The limited Monte Carlo study that we conduct will also give an idea on the relative performance of these methods. We need to define the competing models, the estimation method used for each competing model, and the chi-square type statistic used to measure the departure of each proposed parametric model from the true data generating process. These are now presented.

For our competing models, we consider the problem of choosing between the family of log-normal distributions and the family of exponential distributions. This problem has a long history in the statistical literature. See, e.g., Cox (1962) and Atkinson (1970) among others. The log-normal distribution is parameterized by $\alpha = (\alpha_1, \alpha_2)$ and has density

$$f(x;\alpha_1,\alpha_2) = \frac{1}{x(2\pi)^{1/2}\alpha_2} \exp\left(-\frac{(\log x-\alpha_1)^2}{2\alpha_2^2}\right) \qquad \text{for } x > 0$$

and zero otherwise. The exponential distribution with parameter β has density

$$g(x;\beta) = \frac{1}{\beta} \exp(-x/\beta) \quad \text{for } x > 0$$

and zero otherwise.

The estimator used for each competing model is the ungrouped maximum likelihood estimator (MLE). This choice is particularly convenient here because the ungrouped ML estimator for each model has a closed form and hence is easily computed. Specifically, for log-normal model,

$$\hat{\alpha}_1 = \frac{1}{n} \sum_{i=1}^{n} \log x_i \quad \text{and} \quad \hat{\alpha}_2^2 = \frac{1}{n} \sum_{i=1}^{n} (\log x_i - \hat{\alpha}_1)^2 .$$

For the exponential model, the ungrouped MLE is the sample average, i.e.,

$$\hat{\beta} = \frac{1}{n} \sum_{i=1}^{n} x_i .$$

Lastly, we use the Pearson chi-square statistic to evaluate the discrepancy of a proposed model from the true data generating process. We partition the real line into M intervals $\{(c_{i-1}, c_i),\ i=0,1,\ldots,M\}$ where $c_0 = 0$ and $c_M = +\infty$. The choice of the cells is discussed below. The chi-square statistics for the log-normal and exponential models are:

$$Q_n(\alpha) = n \sum_{i=1}^{M} \frac{(f_i - p_i(\alpha))^2}{p_i(\alpha)}$$

$$Q_n(\beta) = n \sum_{i=1}^{M} \frac{(f_i - p_i(\beta))^2}{p_i(\beta)}$$

where $p_i(\alpha)$ and $p_i(\beta)$ are the probabilities of the interval (c_{i-1}, c_i) under $f(x,\alpha)$ and $g(x,\beta)$, respectively.

In our limited Monte Carlo study, we consider various sets of experiments in which the data are generated from a mixture of an exponential distribution and a log-normal distribution. These two distributions are calibrated so that they have the same population means and variances, namely one and one. Hence the data generating process has the density

$$h(\pi) = \pi \text{ Exponential}(1) + (1-\pi) \text{ Log-normal}(-0.3466, 0.8326),$$

where π is set to some specific value for each set of experiments. In each set of experiments, several random samples are drawn from this mixture of distributions. The sample size varies from 100 to 1000, and for each sample size the number of duplications is 1000.

Throughout, the chosen partition has four cells defined by the values $c_0 = 0$, $c_1 = 0.1$, $c_2 = 1.0$, $c_3 = 3.0$ and $c_4 = +\infty$. Note that because the log-normal distribution has two free parameters, four is the minimum number of cells for which a perfect fit is not always achieved when fitting this distribution by minimum chi-square methods. The power of our tests for model selection is likely to improve by increasing the number of cells. Note also that the shapes of the log-normal and exponential densities differ greatly around the origin. This motivates the choice of $c_1 = 0.1$. The value c_2 is equal to the common population mean, while c_3, which is two standard deviations away from the mean, is used to control for large deviations.

We choose five different values for π which are 0.0, 1.0, 0.5357, 0.25, and 0.75. Although our proposed model selection procedure does not require that the data generating process belong to either of the competing models, we consider the two limiting cases $\pi = 0.0$ and $\pi = 1.0$ for they correspond to the correctly specified cases. The value $\pi = 0.5357$ is determined to be the value for which the estimated log-normal distribution and the estimated exponential distribution are approximately at equal distance from the mixture $h(\pi)$ according to the Pearson discrepancy and the above cells. Thus this set of experiments corresponds approximately to the null hypothesis of our proposed model selection test GCM_n. Finally, to investigate the cases where both competing models are misspecified but not at equal distance from the data generating process, we consider the cases where $\pi = .25$ and $\pi = .75$. The former case corresponds to a data generating process which is log-normal but slightly contaminated by an exponential distribution. The second case is interpreted similarly with an exponential distribution slightly contaminated by a log-normal distribution.

The results of our five sets of experiments are presented in Tables 1-5. The first half of each table gives the average values of the ungrouped ML estimators $\hat{\alpha}$ and $\hat{\beta}$, the Pearson goodness-of-fit statistics $Q_n(\hat{\alpha})$ and $Q_n(\hat{\beta})$, and the model selection statistic GCM_n with its bootstrap estimated variance $\hat{\omega}_*^2$ (see Method 1). The values in parentheses are standard errors. The second half of each table gives in percentage the number of times our proposed model selection procedures based on three methods described in the last section, favor the log-normal model, the exponential model, or are indecisive. The tests are conducted at the 5% nominal significance level. In the first two sets of experiments ($\pi = 0.0$ and $\pi = 1.0$) where one model is correctly specified, we use the labels "correct" and "incorrect" when a choice is made. Finally, in the case where $\pi = .5357$, we give in addition the 2.5%, 5.0%, 95%, and 97.5% fractiles of the observed distribution of the GCM_n statistic. This allows a comparison with the asymptotic N(0,1) approximation under our null hypothesis of equivalence.

Tables 1 and 2 report the cases when one model is correctly specified. It is well-known that the MLE is consistent for the true parameter value under correct specification. For example, in Table 1, the log-normal model is correctly specified, and the MLE of $\alpha=(\alpha_1,\alpha_2)$ approaches the true values $\alpha_o=(-0.3466, 0.8326)$ as the sample size increases from 100 to 1000. However, the Pearson chi-square statistic $Q_n(\hat{\alpha})$ for this model does not have a standard chi-square limiting distribution even under correct specification because the ungrouped MLE is used. In fact, the limiting distribution is somewhere between a $\chi^2(1)$ and a $\chi^2(2)$. See Chernoff and Lehmann (1954). In Table 1, $Q_n(\hat{\alpha})$ has a mean around 1.79 which lies between 1 (mean of $\chi^2(1)$) and 2 (mean of $\chi^2(2)$). For the misspecified model, which is the exponential model here, the MLE $\hat{\beta}$ converges to the pseudo-true parameter β_* which minimizes the KLIC. The corresponding Pearson chi-square statistic $Q_n(\hat{\beta})$, as we expect, increases at the rate of n. The bootstrap estimator of ω also converges as the sample size becomes larger. The test statistic for model selection GCM_n approximately increases at a rate $\sqrt{n}$. In Table 2, where the exponential model is correctly specified, one can observe similar results.

The second half of Table 1 summarizes the results for our three model selection procedures. Method 1 performs quite well and for small sample sizes (n=100 or 250), this method seems to dominate the other two methods in selecting the correct model, which is the log-normal model in this case. However, as the sample size increases to n=500 or 1000, the three methods perform equally well. The three methods also select the correct model almost

100% of the times, as expected.

The second half of Table 2 reports somewhat different results. Except at sample size 1000 where all three methods perform equally well, in smaller samples Method 3 now seems to dominate the other two methods. All three methods, however, do not work as well when the exponential model is correctly specified (Table 2) as when the log-normal model is correctly specified (Table 1). This can be explained by the fact that the log-normal model has one more parameter than the exponential model, and hence is more difficult to reject even when it is misspecified.

For Tables 3, 4 and 5, the data was generated neither from the log-normal model nor from the exponential model, but from a mixture of these two models. Hence, the log-normal model and the exponential model are both incorrectly specified. The MLE's of α and β converge to their pseudo-true values α_o and β_o. For instance, in Table 3, $\alpha=(\alpha_1,\alpha_2)$ converges to $\hat{\alpha}=(-0.4723,\ 1.104)$ and β converges to $\hat{\beta}=0.9994$. The bootstrap estimator $\hat{\omega}_*$ approaches 0.5026. Both chi-square statistics $Q_n(\hat{\alpha})$ and $Q_n(\hat{\beta})$ increase approximately at the rate of n. The same comments apply to Table 4 and Table 5.

In Table 3, the data generating process is chosen such that both the log-normal model and the exponential model are approximately equally close to it. The test statistic GCM_n is expected to have a limiting standard normal distribution N(0,1). This is roughly confirmed in Table 3. For example, for n=1000, GCM_n has a mean of 0.0942 and a standard error of 0.9449. The fractiles reported in Table 3 show that the finite sample distribution of GCM_n is slightly skewed to the right. The three procedures for model selection perform very well. All three of them conclude that both models work equally well in fitting the data with a probability of around 95%, which is 1 minus the nominal size of the test.

With a few exceptions, Tables 4 and 5 reproduce the qualitative results of Tables 1 and 3, respectively, although in a weaker form. When the log-normal model is closer to the true data generating process (Table 4), Method 1 slighlty dominates the other two methods. On the other hand, when the exponential is closer to the true data generating process (Table 5), Method 3 seems to dominate especially at small sample sizes. As noted earlier, selecting the exponential model appears more difficult than selecting the log-normal model.

From our limited Monte Carlo study, it is difficult to say which method absolutely dominates the other methods. Moreover, all methods require about the same amount of computation. Although no clear cut conclusion can be made, our study has shown that the three methods work relatively well.

5. CONCLUSIONS AND FUTURE DIRECTIONS

In this paper, we have studied the problem of selecting estimates using chi-square type statistics and a general class of estimators. In particular, we have proposed some convenient asymptotically standard normal tests based on chi-square type statistics that use estimators in this general class. The tests are designed to determine whether the estimated competing models are as close to the true distribution against the alternative hypothesis that one estimated model is closer, where closeness is measured according to the discrepancy implicit in the chi-square type statistic used.

To facilitate the implementation of our proposed tests, we have used a bootstrap estimate of the asymptotic variance of the numerator of our test statistic. We have also considered two testing procedures that are directly based on the bootstrap method. The three procedures are fairly simple, and mainly require the computation of estimators and chi-square statistics.

Several Monto Carlo experiments were conducted and showed that the three procedures perform relatively well. It was also found that they were comparable, and that none of them absolutely dominates the others.

Our work can be extended in several directions. One direction is to extend our results to econometric models. For econometric models, only the conditional distribution of the endogenous variables y given the exogenous variables z is specified to belong to a conditional parametric probability model $f(y|z;\theta)$, while the marginal distribution of the exogenous variables is left unspecified. Without knowing this marginal distribution, one cannot associate with a given parameter value θ a joint distribution for the observed data (y_i, z_i). Hence when the full sample space $\mathbf{X} = \mathbf{Y} \times \mathbf{Z}$ is partitioned into mutually disjoint cells, the expected probability in each cell cannot be calculated. This expected probability can, however, be consistently estimated by substituting the empirical marginal distribution for the true marginal distribution of z (see Andrews (1988a)). Specifically, we can consider

$$p_i(\theta) = \frac{1}{n} \sum_{j=1}^{n} \int_{Y} 1_{E_i}(y, z_j)\, f(y|z_j;\theta)\, d\nu(y) \quad \text{for } i=1,2,\ldots,M,$$

where $\nu(y)$ is some σ-finite measure on $\mathbf{Y}$. Given these "expected" cell frequencies, chi-square type statistics can be constructed in principle using estimators in the general class considered in this paper. Then the asymptotic standard normality of the resulting test statistics for model selection will likely follow as in section 3.

A second extension is to use random cells instead of fixed cells. Much flexibility will be gained. See, e.g., Andrews (1988b) for various examples. Moreover, with appropriate random cells, the asymptotic distribution of the goodness-of-fit statistics may become independent of the true parameter θ_o under correct specification of the parametric model. See, e.g., Roy (1956) and Watson (1959). Recent work on goodness-of-fit statistics that are asymptotically distributed under correct specification has, however, shown that the asymptotic distribution of such statistics will not change with random cells as long as the random cell boundaries converge in probability to a set of fixed boundaries. See Chibisov (1971), Moore and Spruill (1975) and Andrews (1988a). In view of this latter result, it is expected that our test statistics will remain asymptotically normally distributed with the same asymptotic variance ω^2 under similar conditions.

Appendix

Derivation of ω^2: We first state without proofs two easy lemmas that can also be found in Vuong and Wang (1990).

Lemma 1: Given A3-A4, the weighting matrix $M_n = M(f,\theta_n)$ satisfies

$$M_n = M_o + \frac{1}{\sqrt{n}} L_n + o_p(1/\sqrt{n}),$$

$$L_n = \sum_{i=1}^{M} \frac{\partial M_o}{\partial h_i} \sqrt{n}(f_i - h_i) + \sum_{i=1}^{k} \frac{\partial M_o}{\partial \theta_i} \sqrt{n}(\theta_{ni} - \theta_{oi}),$$

where $M_o = M(h,\theta_o)$, $\partial M_o/\partial h_i$ and $\partial M_o/\partial \theta_i$ are evaluated at (h,θ_o).

Lemma 2: Under A1-A4,

$$\frac{1}{\sqrt{n}} Q_n(\theta_n) = \sqrt{n} b' M_o b + b' L_n b + 2b' M_o D_1 U_n - b' M_o D_2 B \sqrt{n}(\theta_n - \theta_o) + o_p(1).$$

where $b = (\ldots \frac{h_i - p_{oi}}{\sqrt{p_{oi}}}, \ldots)'$, $p_o = p(\theta_o)$ for some $\theta_o \in \Theta$,

$$D_1 = \mathrm{diag}(\ldots, \frac{\sqrt{h_i}}{\sqrt{p_{oi}}}, \ldots),$$

$$D_2 = \mathrm{diag}(\ldots, \frac{h_i + p_{oi}}{p_{oi}} \frac{\sqrt{h_i}}{\sqrt{p_{oi}}}, \ldots) = D_1(D_1^2 + I_M),$$

$B = \mathrm{diag}(\ldots, \frac{1}{\sqrt{h_i}}, \ldots) \ \frac{\partial p_o}{\partial \theta'}$, and $\partial p_o/\partial \theta'$ is evaluated at θ_o,

$$U_n = \sqrt{n}(\ldots, \frac{f_i - h_i}{\sqrt{h_i}}, \ldots)' = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} [e(X_i) - q_H],$$

$$e(X_i) = [1_{E_1}(X_i)/\sqrt{h_1}, \ldots, 1_{E_M}(X_i)/\sqrt{h_M}]' .$$

Both Lemma 1 and Lemma 2 are simply Taylor expansions of M_n at (h,θ_o) and $Q_n(\theta_n)/\sqrt{n}$ at θ_o, respectively. Lemma 2 is also a more detailed expansion of Theorem 5.3 of Moore (1984). However, Moore inadvertently ignored the term $b'L_n b$, and the condition $M_n \dashrightarrow M_*$ in probability under H seems too weak for his stated result.

To obtain the asymptotic variance of $[Q_n(\theta_n)-Q_n(\gamma_n)]/\sqrt{n}$, we define the following 1xM row vectors

$$C'_{1f} = b'_{1f} + 2b'_f M_{fo} D_{1f}, \quad C'_{2f} = (b'_{2f} - b'_f M_{fo} D_{2f} B_f) R^{-1}_{of},$$

$$b'_{1f} = (\ldots, b'_f \frac{\partial M_{fo}}{\partial h_i} b_f, \ldots), \quad b'_{2f} = (\ldots, b'_f \frac{\partial M_{fo}}{\partial \theta_i} b_f, \ldots),$$

a MxM matrix $W_f = E_H[e(X_i)-q_H]\psi(X_i;\theta_o)' = E_H e(X_i)\psi(X_i;\theta_o)'$,

and the other matrices have been defined earlier with a subscript f indicating that the matrices are now attached to the model F_θ. Similar vectors and matrices are defined for the model G_γ with the subscript f replaced by g. Using Lemma 1 and Lemma 2, we can easily obtain

Lemma 3: Given A1-A4,

(i) for model F_θ,

$$\frac{1}{\sqrt{n}} Q_n(\theta_n) = \sqrt{n}Q(\theta_o) + C'_{1f} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} [e(X_i) - q_H] + C'_{2f} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi_f(X_i;\theta_o) + o_p(1),$$

(ii) for model G_γ,

$$\frac{1}{\sqrt{n}} Q_n(\gamma_n) = \sqrt{n}Q(\gamma_o) + C'_{1g} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} [e(X_i) - q_H] + C'_{2g} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \psi_g(X_i;\theta_o) + o_p(1).$$

Note that these results hold under the general H assumed in Section 2. From this lemma it follows that

$$\frac{1}{\sqrt{n}}[Q_n(\theta_n)-Q_n(\gamma_n)] = \sqrt{n}[Q(\theta_o)-Q(\gamma_o)] + (C'_{1f}-C'_{1g},C'_{2f},-C'_{2g})\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\begin{bmatrix} e(X_i)-q_H \\ \psi_f(X_i;\theta_o) \\ \psi_g(X_i;\gamma_o) \end{bmatrix}$$

From the multivariate central limit theorem and assumption A4, we can now immediately obtain the asymptotic distribution of $[Q_n(\theta_n)-Q_n(\gamma_n)]/\sqrt{n}$ under the null hypothesis of equivalence $\tilde{H}_o^e$. Define

$$C'_{fg} = (C'_{1f}-C'_{1g},C'_{2f},-C'_{2g}) \ , \quad W_{fg} = E_H\psi_f(X_i;\theta_o)\psi_g(X_i;\gamma_o)' \ ,$$

$$\Sigma_u = I_M - q_Hq_H' \text{ with } q_H = (\sqrt{h_1},\sqrt{h_2},\ldots,\sqrt{h_M})',$$

$$W_{12} = \begin{bmatrix} \Sigma_u & W_f & W_g \\ W_f' & V_{of} & W_{fg} \\ W_g' & W_{fg}' & V_{og} \end{bmatrix} .$$

Let $\omega^2 = C'_{fg}W_{12}C_{fg}$, we then have $\dfrac{1}{\sqrt{n}}\dfrac{Q_n(\theta_n) - Q_n(\gamma_n)}{\hat{\omega}} \xrightarrow{D} N(0,\omega^2)$

TABLE 1. Data Generating Process = Log-Normal(-0.3466,0.8326).

	100	250	500	1000
$\hat{\alpha}1$	-.3361(.0817)	-.3461(.0543)	-.3471(.0380)	-.3495(.0236)
$\hat{\alpha}2$	.8251(.0559)	.8303(.0393)	.8338(.0251)	.8341(.0203)
$\hat{\beta}$	1.0093(.1016)	.9995(.0655)	1.0019(.0437)	.9984(.0309)
$\hat{\omega}*$	.3495(.2018)	.3349(.1050)	.3149(.0774)	.3076(.050)
$Qn(\hat{\alpha})$	1.730(1.880)	1.740(1.430)	1.740(1.660)	1.790(1.750)
$Qn(\hat{\beta})$	11.79(3.000)	27.32(4.510)	53.29(6.750)	105.15(9.060)
GCMn	-3.707(1.864)	-5.420(2.073)	-7.871(2.416)	-11.028(2.708)

		100	250	500	1000
Model Selection Based On GCM_n	Incorrect Indecisive Correct	0.0% 18.0% 82.0%	0.0% 3.0% 97.0%	0.0% 0.0% 100.0%	0.0% 0.0% 100.0%
Non-bias-corrected Percentile Method	Incorrect Indecisive Correct	0.00% 31.0% 69.0%	0.0% 6.5% 93.5%	0.0% 0.5% 99.5%	0.0% 0.0% 100.0%
Bias-corrected Percentile Method	Incorrect Indecisive Correct	0.00% 42.5% 57.5%	0.0% 8.5% 91.5%	0.0% 0.0% 100.0%	0.0% 0.0% 100.0%

TABLE 2. Data Generating Process = Exponential(1.0)

	100	250	500	1000
$\hat{\alpha}1$	-.5830(.1382)	-.5769(.0753)	-.5790(.0591)	-.5750(.0414)
$\hat{\alpha}2$	1.267(.1270)	1.270(.0854)	1.285(.0588)	1.278(.0406)
$\hat{\beta}$	.9955(.0976)	.9966(.0600)	1.0026(.0440)	1.0010(.0321)
$\hat{\omega}*$	.5614(.1176)	.5240(.0686)	.5140(.0408)	.5070(.0275)
$Qn(\hat{\alpha})$	7.150(4.910)	15.94(7.960)	30.45(11.24)	60.26(16.15)
$Qn(\hat{\beta})$	2.040(1.740)	2.190(2.040)	2.150(1.990)	2.240(2.010)
GCMn	.9171(.8843)	1.6652(.9863)	2.4648(.9820)	3.622(1.003)

		100	250	500	1000
Model Selection Based On GCM_n	Incorrect Indecisive Correct	0.0% 85.5% 14.5%	0.0% 58.5% 41.5%	0.0% 30.5% 69.5%	0.0% 5.0% 95.0%
Non-bias-corrected Percentile Method	Incorrect Indecisive Correct	0.00% 89.0% 11.0%	0.0% 58.0% 42.0%	0.0% 31.5% 68.5%	0.0% 6.5% 93.5%
Bias-corrected Percentile Method	Incorrect Indecisive Correct	0.50% 68.0% 31.5%	0.0% 50.0% 50.0%	0.0% 24.5% 75.5%	0.0% 5.0% 95.0%

TABLE 3. Data Gen. Process=.5357*Exp(1.0)+.4643*log-normal(-0.347,0.833)

		100	250	500	1000
$\hat{\alpha}1$		-.4728(.1107)	-.4743(.0710)	-.4720(.0486)	-.4723(.0328)
$\hat{\alpha}2$		1.097(.1125)	1.102(.0778)	1.106(.0597)	1.104(.0369)
$\hat{\beta}$		.9994(.1027)	.9954(.0640)	1.0003(.0446)	.9994(.0301)
$\hat{\omega}*$		.5016(.0760)	.5040(.0512)	.4986(.0469)	.5026(.0251)
$Qn(\hat{\alpha})$		4.650(3.650)	8.270(5.460)	15.18(7.370)	26.04(9.780)
$Qn(\hat{\beta})$		3.770(2.770)	7.380(4.180)	12.55(5.530)	24.36(7.340)
GCMn		.1109(1.044)	.0737(1.039)	.2100(1.020)	.0942(.9449)
GCM_n	2.5% fractile	-2.895	-2.114	-1.932	-1.811
	5.0% fractile	-1.674	-1.770	-1.643	-1.446
	95.0% fractile	1.6595	1.4666	1.7269	1.5126
	97.5% fractile	2.0570	1.9239	2.0849	2.1395
Model Selection Based On GCM_n	Favor log-n	3.0%	2.5%	1.0%	1.0%
	Equivalent	94.0%	95.0%	94.0%	95.5%
	Favor exp	3.0%	2.5%	5.00%	3.5%
Non-bias-corrected Percentile Method	Favor log-n	3.00%	2.00%	3.50%	1.50%
	Equivalent	96.0%	94.5%	93.0%	95.0%
	Favor exp	1.00%	3.50%	3.50%	3.50%
Bias-corrected Percentile Method	Favor log-n	2.50%	0.50%	1.50%	2.00%
	Equivalent	95.5%	95.0%	95.0%	94.5%
	Favor exp	2.00%	4.50%	3.50%	3.50%

TABLE 4. Data Gen. Process = 0.25*Exp(1.0)+0.75*log-normal(-0.347,0.833)

		100	250	500	1000
$\hat{\alpha}1$		-.4016(.0937)	-.4126(.0586)	-.4078(.0434)	-.4033(.0294)
$\hat{\alpha}2$		.9649(.1105)	.9703(.0602)	.9725(.0486)	.9699(.0345)
$\hat{\beta}$		1.002(.0980)	.9932(.0638)	.9986(.0427)	1.0005(.0289)
$\hat{\omega}*$		.4931(.1403)	.4698(.0541)	.4638(.0372)	.4620(.0310)
$Qn(\hat{\alpha})$		3.130(2.590)	3.960(2.740)	6.030(4.130)	9.970(5.270)
$Qn(\hat{\beta})$		7.320(3.320)	15.63(5.140)	30.14(6.420)	61.40(9.990)
GCMn		-1.126(1.415)	-1.671(1.162)	-2.393(1.053)	-3.589(1.171)
Model Selection Based On GCM_n	Favor log-n	22.0%	33.0%	65.0%	94.0%
	Equivalent	77.5%	67.0%	35.0%	6.0%
	Favor exp	0.5%	0.0%	0.0%	0.0%
Non-bias-corrected Percentile Method	Favor log-n	20.0%	40.5%	64.0%	89.0%
	Equivalent	70.0%	59.5%	36.0%	11.0%
	Favor exp	0.00%	0.00%	0.00%	0.00%
Bias-corrected Percentile Method	Favor log-n	20.0%	28.0%	54.0%	87.5%
	Equivalent	80.0%	72.0%	46.0%	12.5%
	Favor exp	0.00%	0.00%	0.00%	0.00%

TABLE 5. Data Gen. Process = 0.75*Exp(1.0)+0.25*log-normal(-0.347,0.833)

		100	250	500	1000
$\hat{\alpha}1$		-.5220(.1158)	-.5171(.0715)	-.5323(.0585)	-.5215(.0375)
$\hat{\alpha}2$		1.171(.1188)	1.181(.0823)	1.194(.0607)	1.193(.0399)
$\hat{\beta}$		.9917(.1015)	.9984(.0618)	.9894(.0466)	.9982(.0309)
$\hat{\omega}*$		.5227(.1184)	.5108(.0548)	.4969(.0351)	.5023(.0235)
$Qn(\hat{\alpha})$		5.740(4.240)	11.37(6.690)	22.53(8.770)	41.73(12.60)
$Qn(\hat{\beta})$		2.620(2.410)	3.920(2.980)	4.910(3.550)	8.510(4.860)
GCMn		.5482(1.003)	.9013(.9879)	1.5761(.9453)	2.0816(.9309)
Model Selection Based On GCM_n	Favor log-n	0.20%	0.05%	0.00%	0.0%
	Equivalent	93.5%	88.0%	62.5%	43.0%
	Favor exp	4.50%	11.5%	37.5%	57.0%
Non-bias-corrected Percentile Method	Favor log-n	2.00%	0.50%	0.00%	0.00%
	Equivalent	93.5%	81.0%	75.0%	39.5%
	Favor exp	4.50%	18.5%	25.0%	60.5%
Bias-corrected Percentile Method	Favor log-n	1.00%	0.00%	0.00%	0.00%
	Equivalent	89.0%	80.0%	66.0%	37.0%
	Favor exp	10.0%	20.0%	34.0%	63.0%

REFERENCES

Akaike, H.(1973):"Information Theory and an Extension of the Likelihood Ratio Principle," *Proceedings of the Second International Symposium of Information Theory*, ed. by B.N.Petrov and F.Csaki. Budapest:Akademiai Kiado, 257-281.

Amemiya, T.(1985): *Advanced Econometrics*. Cambridge: Harvard University Press.

Andrews, D.W.K.(1988a):"Chi-Square Diagnostic Tests for Econometric Models: Theory," *Econometrica*, 56, 1419-1453.

------(1988b):"Chi-Square Diagnostic Tests for Econometric Models: Introduction and Applications," *Journal of Econometrics*, 37, 135-156.

Atkinson, A.C.(1970):"A Method for Discriminating Between Models," *Journal of the Royal Statistical Society, Series B*, 32, 323-353.

Bishop, Y.M.M., S.E. Fienberg, and P.W. Holland (1975): *Discrete Multivariate Analysis*. Cambridge: MIT Press.

Chernoff, H. and E.L. Lehmann(1954):"The Use of Maximum of Likelihood Estimates in χ^2 Test for Goodness of Fit," *Ann.Math.Statist.*, 25, 579-586.

Chibisov, D.M.(1971):"Certain Chi-Square Type Tests for Continuous Distributions," *Theory of Probability and Its Applications*,16, 1-22.

Cochran, W.G. (1952):"The χ^2 Test of Goodness of Fit," *Ann. Math. Statist.* 23, 315-345.

Cox, D.R.(1962):"Further Results on Tests of Separate Families of Hypotheses," *Journal of the Royal Statistical Society*, Series B, 24, 406-424.

Efron, B. (1982):"The Jackknife, the Bootstrap and Other Resampling Plans," *CBMS-NSF Regional Conference Series in Applied Mathematics*, 38 .

Efron, B. (1984): "Comparing Non-Nested linear Models," *Journal of the American Statistical Association*, 79, 791-803.

Heckman, J.J.(1981):"Heterogeneity and State Dependence," in *Studies in Labor Markets*, (University of Chicago Press, Chicago) ed. by S. Sosen.

Linhart, H. and Zucchini, W.(1986): *Model Selection*, John Wiley and Sons.

Massy, W.F., D.B. Montgomery and D.G. Morrison (1970):*Stochastic Models of Buying Behavior*, Cambridge, Massachusetts: MIT Press.

Moore, D.S.(1977):"Generalized Inverses, Wald's Method and the Construction of Chi-Squared Tests of Fit," *Journal of the American Statistical Association*, 7, 131-137.

------(1978):"Chi-Square Tests," in *Studies in Statistics*, ed. by R.V. Hogg. Volume 19, The Mathematical Association of America.

----- (1984): "Measures of Lack of Fit from Tests of Chi-Squared Type," *Journal of Statistical Planning and Inference*, 10, 151-166.

-----(1986): "Tests of Chi-squared Type," *Goodness-of-Fit techniques*, ed. by R.B. D'Agostino and M.A. Stephens.

Moore, D.S. and M.C. Spruill(1975):"Unified Large-Sample theory of General Chi-squared Statistics for Tests of Fit," *Ann. Statist*. 3, 599-616.

Nakamura, A. and M. Nakamura (1985):*The Second Paycheck: A Socioeconomic Analysis of Earnings*, Orlando, Florida: Academmic Press.

Pearson, K. (1900):"On the Criterion that a Given System of Deviation from the Probable in the Case of a Correlated System of Variables is such that It Can be reasonablly supposed to have arisen from Random sampling," *The London, edinburgh and Dublin Philosophical Magazine and Journal of Science*, 50, 157-175.

Rao, K.C. and D.S. Robson(1974):"A Chi-Square Statistic for Goodness-of-Fit Tests within the exponential family," *Communications in Statistics*, 3, 1139-1153.

Roy, A.R. (1956):"On χ^2 Statistics with Variable Intervals," Technical Report No.1, Stanford University, Department of Statistics.

Vuong, Q.H.(1989): "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses," *Econometrica*, 57. 257-306.

Vuong, Q.H., and W. Wang(1990):"Minimum Chi-Square Estimation and Tests for Model Selection," MRG Working Paper No.M9003, University of Southern California.

Watson, G.S.(1959):"Some Recent Results in Chi-Square Goodness-of-Fit Tests," *Biometrics*, 15, 440-468.

White, H.(1982): "Maximum Likelihood Estimation of Misspecified Models," *Econometrica*, 50. 1-25.

Discussion Paper Series, CentER, Tilburg University, The Netherlands:

(For previous papers please consult previous discussion papers.)

No.	Author(s)	Title
9051	M. Baye, D. Kovenock and C. de Vries	The All-Pay Auction with Complete Information
9052	H. Carlsson and E. van Damme	Global Games and Equilibrium Selection
9053	M. Baye and D. Kovenock	How to Sell a Pickup Truck: "Beat-or-Pay" Advertisements as Facilitating Devices
9054	Th. van de Klundert	The Ultimate Consequences of the New Growth Theory; An Introduction to the Views of M. Fitzgerald Scott
9055	P. Kooreman	Nonparametric Bounds on the Regression Coefficients when an Explanatory Variable is Categorized
9056	R. Bartels and D.G. Fiebig	Integrating Direct Metering and Conditional Demand Analysis for Estimating End-Use Loads
9057	M.R. Veall and K.F. Zimmermann	Evaluating Pseudo-R^2's for Binary Probit Models
9058	R. Bartels and D.G. Fiebig	More on the Grouped Heteroskedasticity Model
9059	F. van der Ploeg	Channels of International Policy Transmission
9060	H. Bester	The Role of Collateral in a Model of Debt Renegotiation
9061	F. van der Ploeg	Macroeconomic Policy Coordination during the Various Phases of Economic and Monetary Integration in Europe
9062	E. Bennett and E. van Damme	Demand Commitment Bargaining: - The Case of Apex Games
9063	S. Chib, J. Osiewalski and M. Steel	Regression Models under Competing Covariance Matrices: A Bayesian Perspective
9064	M. Verbeek and Th. Nijman	Can Cohort Data Be Treated as Genuine Panel Data?
9065	F. van der Ploeg and A. de Zeeuw	International Aspects of Pollution Control
9066	F.C. Drost and Th. E. Nijman	Temporal Aggregation of GARCH Processes
9067	Y. Dai and D. Talman	Linear Stationary Point Problems on Unbounded Polyhedra
9068	Th. Nijman and R. Beetsma	Empirical Tests of a Simple Pricing Model for Sugar Futures

No.	Author(s)	Title
9069	F. van der Ploeg	Short-Sighted Politicians and Erosion of Government Assets
9070	E. van Damme	Fair Division under Asymmetric Information
9071	J. Eichberger, H. Haller and F. Milne	Naive Bayesian Learning in 2 x 2 Matrix Games
9072	G. Alogoskoufis and F. van der Ploeg	Endogenous Growth and Overlapping Generations
9073	K.C. Fung	Strategic Industrial Policy for Cournot and Bertrand Oligopoly: Management-Labor Cooperation as a Possible Solution to the Market Structure Dilemma
9101	A. van Soest	Minimum Wages, Earnings and Employment
9102	A. Barten and M. McAleer	Comparing the Empirical Performance of Alternative Demand Systems
9103	A. Weber	EMS Credibility
9104	G. Alogoskoufis and F. van der Ploeg	Debts, Deficits and Growth in Interdependent Economies
9105	R.M.W.J. Beetsma	Bands and Statistical Properties of EMS Exchange Rates
9106	C.N. Teulings	The Diverging Effects of the Business Cycle on the Expected Duration of Job Search
9107	E. van Damme	Refinements of Nash Equilibrium
9108	E. van Damme	Equilibrium Selection in 2 x 2 Games
9109	G. Alogoskoufis and F. van der Ploeg	Money and Growth Revisited
9110	L. Samuelson	Dominated Strategies and Commom Knowledge
9111	F. van der Ploeg and Th. van de Klundert	Political Trade-off between Growth and Government Consumption
9112	Th. Nijman, F. Palm and C. Wolff	Premia in Forward Foreign Exchange as Unobserved Components
9113	H. Bester	Bargaining vs. Price Competition in a Market with Quality Uncertainty
9114	R.P. Gilles, G. Owen and R. van den Brink	Games with Permission Structures: The Conjunctive Approach
9115	F. van der Ploeg	Unanticipated Inflation and Government Finance: The Case for an Independent Common Central Bank

No.	Author(s)	Title
9116	N. Rankin	Exchange Rate Risk and Imperfect Capital Mobility in an Optimising Model
9117	E. Bomhoff	Currency Convertibility: When and How? A Contribution to the Bulgarian Debate!
9118	E. Bomhoff	Stability of Velocity in the G-7 Countries: A Kalman Filter Approach
9119	J. Osiewalski and M. Steel	Bayesian Marginal Equivalence of Elliptical Regression Models
9120	S. Bhattacharya, J. Glazer and D. Sappington	Licensing and the Sharing of Knowledge in Research Joint Ventures
9121	J.W. Friedman and L. Samuelson	An Extension of the "Folk Theorem" with Continuous Reaction Functions
9122	S. Chib, J. Osiewalski and M. Steel	A Bayesian Note on Competing Correlation Structures in the Dynamic Linear Regression Model
9123	Th. van de Klundert and L. Meijdam	Endogenous Growth and Income Distribution
9124	S. Bhattacharya	Banking Theory: The Main Ideas
9125	J. Thomas	Non-Computable Rational Expectations Equilibria
9126	J. Thomas and T. Worrall	Foreign Direct Investment and the Risk of Expropriation
9127	T. Gao, A.J.J. Talman and Z. Wang	Modification of the Kojima-Nishino-Arima Algorithm and its Computational Complexity
9128	S. Altug and R.A. Miller	Human Capital, Aggregate Shocks and Panel Data Estimation
9129	H. Keuzenkamp and A.P. Barten	Rejection without Falsification - On the History of Testing the Homogeneity Condition in the Theory of Consumer Demand
9130	G. Mailath, L. Samuelson and J. Swinkels	Extensive Form Reasoning in Normal Form Games
9131	K. Binmore and L. Samuelson	Evolutionary Stability in Repeated Games Played by Finite Automata
9132	L. Samuelson and J. Zhang	Evolutionary Stability in Asymmetric Games
9133	J. Greenberg and S. Weber	Stable Coalition Structures with Uni-dimensional Set of Alternatives

No.	Author(s)	Title
9134	F. de Jong and F. van der Ploeg	Seigniorage, Taxes, Government Debt and the EMS
9135	E. Bomhoff	Between Price Reform and Privatization - Eastern Europe in Transition
9136	H. Bester and E. Petrakis	The Incentives for Cost Reduction in a Differentiated Industry
9137	L. Mirman, L. Samuelson and E. Schlee	Strategic Information Manipulation in Duopolies
9138	C. Dang	The D_2^*-Triangulation for Continuous Deformation Algorithms to Compute Solutions of Nonlinear Equations
9139	A. de Zeeuw	Comment on "Nash and Stackelberg Solutions in a Differential Game Model of Capitalism"
9140	B. Lockwood	Border Controls and Tax Competition in a Customs Union
9141	C. Fershtman and A. de Zeeuw	Capital Accumulation and Entry Deterrence: A Clarifying Note
9142	J.D. Angrist and G.W. Imbens	Sources of Identifying Information in Evaluation Models
9143	A.K. Bera and A. Ullah	Rao's Score Test in Econometrics
9144	B. Melenberg and A. van Soest	Parametric and Semi-Parametric Modelling of Vacation Expenditures
9145	G. Imbens and T. Lancaster	Efficient Estimation and Stratified Sampling
9146	Th. van de Klundert and S. Smulders	Reconstructing Growth Theory: A Survey
9147	J. Greenberg	On the Sensitivity of Von Neuman and Morgenstern Abstract Stable Sets: The Stable and the Individual Stable Bargaining Set
9148	S. van Wijnbergen	Trade Reform, Policy Uncertainty and the Current Account: A Non-Expected Utility Approach
9149	S. van Wijnbergen	Intertemporal Speculation, Shortages and the Political Economy of Price Reform
9150	G. Koop and M.F.J. Steel	A Decision Theoretic Analysis of the Unit Root Hypothesis Using Mixtures of Elliptical Models
9151	A.P. Barten	Consumer Allocation Models: Choice of Functional Form

No.	Author(s)	Title
9152	R.T. Baillie, T. Bollerslev and M.R. Redfearn	Bear Squeezes, Volatility Spillovers and Speculative Attacks in the Hyperinflation 1920s Foreign Exchange
9153	M.F.J. Steel	Bayesian Inference in Time Series
9154	A.K. Bera and S. Lee	Information Matrix Test, Parameter Heterogeneity and ARCH: A Synthesis
9155	F. de Jong	A Univariate Analysis of EMS Exchange Rates Using a Target Zone Model
9156	B. le Blanc	Economies in Transition
9157	A.J.J. Talman	Intersection Theorems on the Unit Simplex and the Simplotope
9158	H. Bester	A Model of Price Advertising and Sales
9159	A. Özcam, G. Judge, A. Bera and T. Yancey	The Risk Properties of a Pre-Test Estimator for Zellner's Seemingly Unrelated Regression Model
9160	R.M.W.J. Beetsma	Bands and Statistical Properties of EMS Exchange Rates: A Monte Carlo Investigation of Three Target Zone Models
9161	A.M. Lejour and H.A.A. Verbon	Centralized and Decentralized Decision Making on Social Insurance in an Integrated Market
9162	S. Bhattacharya	Sovereign Debt, Creditor-Country Governments, and Multilateral Institutions
9163	H. Bester, A. de Palma, W. Leininger, E.-L. von Thadden and J. Thomas	The Missing Equilibria in Hotelling's Location Game
9164	J. Greenberg	The Stable Value
9165	Q.H. Vuong and W. Wang	Selecting Estimated Models Using Chi-Square Statistics

P.O. BOX 90153. 5000 LE TILBURG. THE NETHERLANDS