Journal of Statistical Planning and Inference 97 (2001) 85–100

journal of statistical planning and inference

www.elsevier.com/locate/jspi

Rao's score test with nonparametric density estimators

Gloria González-Rivera, Aman Ullah *

Department of Economics, University of California- Riverside, Riverside, CA 92521, USA

Abstract

Traditionally, Rao's score (RS) tests are constructed under a parametric specification of the probability density function. We estimate the density function by a non-parametric estimator and consider a semi-parametric Rao's score (SPRS) test for a set of hypotheses concerning the parametric model. The asymptotic distribution of the SPRS test is analyzed. Further, for the regression model, we carry out a set of Monte Carlo experiments to analyze the size and power of the SPRS test in small samples. The robustness of SPRS test to the choice of the density estimator is also analyzed. © 2001 Elsevier Science B.V. All rights reserved.

MSC: 62G10; 62G20

Keywords: Score test; Semiparametric; Kernel density; Regression

1. Introduction

The well-known Rao's score (RS) test is commonly used in econometrics for testing restrictions on the parameters of the economic models which have prespecified parametric components. For example, testing the regression coefficient to be zero in a model where the systematic component is linear and the error component has a known parametric probability density, usually normal; see Godfrey (1988), Bera and Ullah (1991) and Bera and Bilias (1997) for more examples. The RS test, called in this paper the parametric RS (PRS) test, explores the distance, from zero, of the sample score function evaluated at the constrained maximum-likelihood (ML) estimator of the parameters under the null hypothesis. The test statistic is defined as the weighted sum of squares of the sample score function where the weight is the inverse of the variance of the sample score. Since the sample score evaluated at the true parameter value is asymptotically normal with zero mean the test statistic is asymptotically distributed as a central χ^2 distribution.

* Corresponding author.

E-mail addresses: ggonzale@galaxy.ucr.edu (G. González-Rivera), ullah@mail.ucr.edu (A. Ullah).

PII: S0378-3758(00)00347-5

There are many instances in which the parametric probability density may be specified as a non-normal density. Under this scenario, one can still consider the misspecified normal probability density and construct the RS test based on the quasi-ML estimators (QMLE). Such a test will be called the QPRS test. However the QMLE, though consistent, may be quite inefficient and hence may affect the power property of the test. In view of above we consider here the RS tests which do not depend on the parametric specification of the probability density. We only assume that the density function is sufficiently smooth to be estimated by a consistent nonparametric estimator. Such a RS test is referred to as the semiparametric RS (SPRS) test.

The idea of nonparametric kernel estimation of unknown density and hence the score to develop the efficient estimation and testing was originated in the work of Stone (1975), Bickel (1978, 1982) and Begun et al. (1983); also see Kreiss (1990), Newey (1994), Newey and McFadden (1994), Pagan and Ullah (1996, Chapter 5), and the interesting work of Silvapulle et al. (1997). The works of Pagan and Ullah (1996) and Silvapulle et al. (1997) follow the adaptive estimation literature (Bickel, 1982; Bickel et al., 1993) to develop score tests. In an important recent paper Ai (1997) studied the asymptotic normality of the sample score function and the ML estimator for a class of SP models with nonparametric kernel density estimators. This result is a key ingredient for our proof of the SPRS test to be asymptotically distributed as a central χ^2, and in this respect our results differ with the previous studies.

Our proposed SPRS test can be used for a general class of SP economic models where some components are parametric and the others are nonparametric. This includes, for example, the SP sample selection models, and the models where the systematic component is partially specified as in Robinson (1988) and Ichimura and Lee (1991). Further, we show here that the SPRS tests can be computed as the nR^2, where n is the number of observations and R^2 is the uncentered coefficient of determination in the regression of ones on the score. An additional contribution of the paper is to conduct Monte Carlo experiments to analyze the performance of the size and power of the SPRS test in small samples. It has been found that there is a positive pay-off to the use of the SPRS tests. For the problem of testing zero restrictions on the regression parameters, the SPRS tests performed as well as the QPRS test. But, for the problem of heteroskedasticity, the SPRS tests out performed the QPRS test. Finally, we also analyze the robustness of test to the choice of the nonparametric density estimator. The SPRS test is consistent and it is asymptotically distributed as a central χ^2 distribution under the null for a consistent estimator of the density. However, this may not be true in small samples. To analyze the robustness question we considered the traditional Rosenblatt's (1956) kernel estimator and a recently developed kernel estimator with a parametric start due to Hjort and Glad (1995), which is claimed to be better than the traditional estimator in the mean-squared error (MSE) sense. In the hypothesis testing contexts, considered here, our results suggests that the traditional kernel estimator performs better than the Hjort and Glad (1995) estimator in the power sense.

The plan of the paper is as follows. In Section 2 we present the model and the SPRS test. Then in Section 3 we consider the applications of the SPRS test. Finally, in Section 4 we present the Monte Carlo results.

2. SPRS tests

Let us consider the economic variables under study as $z=(y,x)$ where y is a scalar random variable and x is a random vector.[1] Then a SP econometric model relating y to x can be given by the conditional density of y given x, $\phi(y\,|\,x,\theta,f)$, where θ represents the p dimensional column of parameters, and f represents the infinite dimensional parameters. Following Ai (1997) we assume that $\phi(\cdot)$ satisfies an index restriction[2]

$$\phi(y\,|\,x,\theta,f)=J(z,\theta)f(v_1(z,\theta)\,|\,v_2(x,\theta),\theta), \tag{2.1}$$

where $f()$ is the conditional density of $v_1(z,\theta)$ given $v_2(x,\theta)$ and $J(z,\theta)$ is the known Jacobian of the transformation from $v_1(z,\theta)$ to y. The focus here will only be on the SP models where the conditional density $f()$ is unknown. For example, the SP nonlinear models

$$y=m(x,\theta)+u, \tag{2.2}$$

where m is a known function but the random error u has an unknown density $f(u)$. For this model $v_1(z,\theta)=y-m(x,\theta)$, $v_2(x,\theta)=x$ and $J(z,\theta)=1$ so that

$$\phi(y\,|\,x,\theta,f)=f((y-m(x,\theta))\,|\,x)=f(u\,|\,x), \tag{2.3}$$

where $f(u\,|\,x)=f(u)$ if x and u are independent. Ai (1997) proposed a SPML estimator of θ for the SP models given by (2.1). If $f()$ is assumed to have a known parametric form then one can use the parametric ML estimation of θ.

The score function of (2.1) or (2.3) is denoted by

$$s(z,\theta,f)=\frac{\partial}{\partial\theta}\log\phi(y\,|\,x,\theta,f). \tag{2.4}$$

Further define θ_0 as the true value of the parameter θ. Then it follows that $E[s(z,\theta,f)]=0$ if and only if $\theta=\theta_0$. Let $z_i=\{y_i,x_i\}$, $i=1,\dots,n$, represents an i.i.d. sample of size n. Then the log-likelihood function is

$$\ell(\theta)=\sum_1^n \log\phi(y_i\,|\,x_i,\theta,f) \tag{2.5}$$

and the sample score function is

$$s(\theta)=\frac{\partial\ell(\theta)}{\partial\theta}=\sum_1^n s(z_i,\theta,f)=S'(\theta)\imath, \tag{2.6}$$

[1] The results of this paper also hold for the case where y is a random vector.

[2] Note that this index restriction is not restrictive in the sense that many statistical and econometric models satisfy this restriction, for example, partially linear regression model, sample selection model, simultaneous equations model and duration model; see Ai (1997) for more details.

where $S(\theta)$ is an $n \times p$ matrix of score vectors $s(z_i, \theta, f)'$ and $\imath$ is an $n \times 1$ vector of ones. $E(s(\theta)) = 0$ if and only if $\theta = \theta_0$. Further, defining $\Sigma = E[s(z_i, \theta, f)\, s(z_i, \theta, f)']$,

$$V\left(\frac{s(\theta)}{\sqrt{n}}\right) = \frac{1}{n} E[s(\theta)s(\theta)'] = E[s(z_i, \theta, f)s(z_i, \theta, f)'] = \Sigma. \tag{2.7}$$

We note that if $f = f_0$ is the true density and $\theta = \theta_0$, $V(s(\theta)) = I(\theta) = -E(\partial^2 \ell(\theta)/\partial\theta\partial\theta') = -E((\partial/\partial\theta)\sum_1^n s(z_i, \theta, f)')$; $I(\theta)$ is the information matrix.

We consider the problem of testing

$$\mathrm{H}_0 \colon g(\theta) = 0 \quad \text{against } \mathrm{H}_1 \colon g(\theta) \neq 0, \tag{2.8}$$

where $g(\theta)$ is a $q \times 1$ vector of known functions with $G(\theta) = \partial g(\theta)/\partial\theta'$. For this testing problem Newey and McFadden (1994) considered the SPRS test as[3]

$$\begin{aligned} \mathrm{SPRS}^* &= s(\tilde{\theta})'(V(s(\tilde{\theta})))^{-1} s(\tilde{\theta}) \\ &= \frac{1}{n} s(\tilde{\theta})' \tilde{\Sigma}^{-1} s(\tilde{\theta}), \end{aligned} \tag{2.9}$$

where $\tilde{\theta}$ is the ML estimator, under the null, for the SP models with known f, $V(s(\tilde{\theta}))$ is the consistent estimate of the asymptotic covariance matrix $V(s(\theta)) = n\Sigma$ evaluated at $\theta = \tilde{\theta}$ and a consistent estimator of Σ is

$$\tilde{\Sigma} = \frac{1}{n} \sum_1^n s(z_i, \tilde{\theta}, f)s(z_i, \tilde{\theta}, f)' = \frac{1}{n} S'(\tilde{\theta})S(\tilde{\theta}). \tag{2.10}$$

We note that the second equality in (2.9) holds up to $\mathrm{o_p}(1)$.

The ML estimator $\tilde{\theta}$, under the null $g(\theta) = 0$, is obtained from the first-order conditions of the Lagrangean function $\ell(\theta) + \lambda' g(\theta)$ with respect to θ and the Lagrange multiplier parameter λ. This is the solution of $s(\tilde{\theta}) + G'(\tilde{\theta})\tilde{\lambda} = 0$ and $g(\tilde{\theta}) = 0$. Using linearization the iterative solution, under the null of $g(\theta) = 0$, is given by

$$\begin{pmatrix} \tilde{\theta} \\ \tilde{\lambda} \end{pmatrix} \simeq \begin{pmatrix} \theta \\ 0 \end{pmatrix} + \begin{bmatrix} \dfrac{-\partial s(\theta)}{\partial\theta'} & G'(\theta) \\ G(\theta) & 0 \end{bmatrix}^{-1} \begin{bmatrix} s(\theta) \\ g(\theta) \end{bmatrix}. \tag{2.11}$$

Newey and McFadden (1994) showed that, under the null, the asymptotic distribution of (2.9) is a central χ^2_q which is the same as the asymptotic distributions of the PRS and QPRS tests.

The SPRS^* test in (2.9) is, however, not operational. This is because both the score function $s(\theta)$ in (2.6) and the variance function Σ in (2.7) depend upon the unknown f. To obtain an operational SPRS test we first consider a consistent nonparametric estimator of f. Here we consider two such estimators. First is the traditional Rosenblatt's (1956) kernel density estimator

$$\hat{f}(u) = \frac{1}{n} \sum_1^n K(u_i - u), \tag{2.12}$$

[3] Newey and McFadden (1994) refer to this the Lagrange multiplier (LM) test. See Bera and Ullah (1993) and Bera and Bilias (1997) on this point. Since the parameter θ may belong to a SP model given by (2.1) the RS test in (2.9) is referred to as the SPRS test.

where $K(u_i - u) = h^{-1}K_0((u_i - u)/h)$, u_i, $i = 1 \ldots, n$, are the i.i.d. observations, K_0 is the kernel function and h is the window width. For our simulation results in Section 4 we use normal kernel and $h \propto n^{-1/5}$. For the details on the choice of h and K_0, and the asymptotic properties of $\hat{f}$, see Silverman (1986), Marron (1988) and Pagan and Ullah (1996).

The second consistent estimator considered is due to Hjort and Glad (1995), and it is given by

$$\hat{f}(u) = f(u, \eta)\hat{r}(u), \tag{2.13}$$

where $f(u, \eta)$ is the parametric start function, for instance a normal density with mean μ and variance σ^2, and the remainder

$$\hat{r}(u) = \frac{1}{n}\sum_{1}^{n} \frac{K(u_i - u)}{f(u_i, \eta)}. \tag{2.14}$$

The role of the remainder is to adjust for deviations of the assumed parametric start density from the true density function. Hjort and Glad (1995) demonstrate that the estimator (2.13) performs better than the traditional kernel estimator in the MSE sense.

An operational SPRS test can now be proposed as

$$\begin{aligned} \text{SPRS} &= \hat{s}(\hat{\theta})'(V(\hat{s}(\hat{\theta})))^{-1}\hat{s}(\hat{\theta}) \\ &= \frac{1}{n}\hat{s}(\hat{\theta})'\hat{\Sigma}^{-1}\hat{s}(\hat{\theta}) \\ &= \iota'\hat{S}(\hat{\theta})(\hat{S}'(\hat{\theta})\hat{S}(\hat{\theta}))^{-1}\hat{S}'(\hat{\theta})\iota = nR^2, \end{aligned} \tag{2.15}$$

where $\hat{s}(\hat{\theta}) = \sum_1^n s(z_i, \hat{\theta}, \hat{f}) = \hat{S}(\hat{\theta})'\iota, \hat{\theta}$ is the consistent SPML estimator under the null with f replaced by $\hat{f}$,[4,5]

$$\hat{\Sigma} = \frac{1}{n}\sum_{1}^{n} s(z_i, \hat{\theta}, \hat{f})s(z_i, \hat{\theta}, \hat{f})' = \frac{1}{n}\hat{S}'(\hat{\theta})\hat{S}(\hat{\theta}) \tag{2.16}$$

and R^2 is the uncentered coefficient of determination of the regression of ones on the scores $s(z_i, \hat{\theta}, \hat{f})$. The third and fourth equalities in (2.15) hold asymptotically. The SPML estimator $\hat{\theta}$ is the solution of $\hat{s}(\hat{\theta}) + G'(\hat{\theta})\hat{\lambda} = 0$ and $g(\hat{\theta}) = 0$. Using linearization, $\hat{\theta}$ is given by the iterative solution under the null of $g(\theta) = 0$

$$\begin{pmatrix} \hat{\theta} \\ \hat{\lambda} \end{pmatrix} \simeq \begin{pmatrix} \theta \\ 0 \end{pmatrix} + \begin{bmatrix} \dfrac{-\partial\hat{s}(\theta)}{\partial\theta'} & G'(\theta) \\ G(\theta) & 0 \end{bmatrix}^{-1} \begin{bmatrix} \hat{s}(\theta) \\ g(\theta) \end{bmatrix}. \tag{2.17}$$

In practice one can replace the initial value of θ in (2.17) by a consistent estimator θ^* which gives $\hat{\theta} = \hat{\theta}^*$ and $\hat{\lambda} = \hat{\lambda}^*$ as a usual one-step estimator. The estimator $\hat{\theta}^*$, under the null, is asymptotically equivalent to the iterative estimator $\hat{\theta}$, see Appendix. This

[4] Although it is not studied here, a Wald-type test can be developed as $W = g(\hat{\theta}_1)'\hat{\Omega}^{-1}g(\hat{\theta}_1)$ where $\hat{\Omega}$ is a consistent estimator of the asymptotic variance, Ω, where $\Omega = V(g(\hat{\theta}_1)) = \partial g(\theta)/\partial\theta'\Sigma^{-1}\partial g(\theta)/\partial\theta$ evaluated at the SPML estimator $\hat{\theta}_1$ under the alternative.

[5] Ai (1997) suggests trimming of $s(z_i, \hat{\theta}, \hat{f})$ to avoid small density values and the boundary bias.

implies that the one-step estimator $\hat{\theta}^*$ can be used in place of $\hat{\theta}$ to develop SPRS test statistics.

In the appendix we show that, under the null, the SPRS test (2.15) is asymptotically distributed as a central χ^2_q. Thus the asymptotic distribution of the SPRS test is the same as the PRS test. However this may not be true in small samples which is one of the objectives of the Monte Carlo simulations in Section 4.

3. Applications

The SPRS test considered in Section 2 can be used for various SP models described by (2.1). Here we give some examples for testing restrictions in a linear model with the unknown density of errors.

3.1. Testing restrictions in a linear model

Consider a SP regression model

$$y_i = x_i'\beta + u_i, \tag{3.1}$$

where x_i is a $l \times 1$ vector of regressors written in deviation from their means, u_i are i.i.d. error terms with unknown form of the probability density function $f(u)$, and x and u are assumed to be independent. We partition $x_i' = (x_{1i}'\ x_{2i}')$ and $\beta = (\beta_1'\beta_2')'$ where x_{1i} and β_1 are each $\ell_1 \times 1$ and x_{2i} and β_2 are each $\ell_2 \times 1$ sub vectors, respectively, $\ell_1 + \ell_2 = \ell$. Our interest is to test a set of restrictions on the subvector of parameters β_1

$$\mathrm{H}_0: \beta_1 = \beta_1^0, \qquad \mathrm{H}_1: \beta_1 \neq \beta_1^0.$$

Model (3.1) satisfies (2.1) with $\theta = \beta$, $v_1(z_i, \theta) = y_i - x_i'\beta = u_i, v_2(x_i, \theta) = x_i$ and $J(z_i, \theta) = 1$.

The score function for the SP model (3.1) is

$$s(\beta) = \sum_1^n s(z_i, \beta, f) = -\sum_1^n x_i\psi(u_i) = -X'\psi, \tag{3.2}$$

where $\psi(u_i) = \psi_i = f^{(1)}(u_i)/f(u_i)$, $f^{(1)}$ is the first derivative of f with respect to u_i, ψ is an $n \times 1$ vector, and X is an $n \times l$ matrix. For the true value of the parameter vector $Es(\beta) = 0$ and using (2.7) and (3.2).

$$V(s(\beta)) = E(s(\beta)s(\beta)') = E(X'\psi\ \psi'X) = n\Sigma. \tag{3.3}$$

Consequently, a consistent estimator of Σ is

$$\hat{\Sigma} = \left(\frac{\hat{\psi}'\hat{\psi}}{n}\right)\left(\frac{X'X}{n}\right) \tag{3.4}$$

and the SPRS test (2.15) becomes

$$\mathrm{SPRS} = n\frac{\hat{\psi}'X(X'X)^{-1}X'\hat{\psi}}{\hat{\psi}'\hat{\psi}} = \iota'\hat{S}(\hat{S}'\hat{S})^{-1}\hat{S}'\iota = nR^2, \tag{3.5}$$

where $\hat{\psi}_i = \psi(y_i - x_i'\hat{\beta})$ is ψ_i evaluated under the null, $\hat{\beta}$ is a column vector of β_1^0 and $\hat{\beta}_2$, and R^2 is the uncentered coefficient of determination of the regression of $\hat{\psi}$ on X or, alternatively, the regression of the vector ι of ones on the matrix $\hat{S} = S(\hat{\beta})$ of scores $-x_i\hat{\psi}_i$. The second and third equalities in (3.5) hold asymptotically. Since the regressors in (3.1) are specified in deviations from their means the score $s(\beta)$ in our model coincides with the efficient score in Begun et al. (1983) sense,[6] see Ai (1997) and Pagan and Ullah (1996) among others. This implies that the SPRS test in (3.5) is the locally most powerfully test, see also Silvapulle et al. (1997).

For the calculations of $\hat{\psi}$ we note that, under the traditional Rosenblatt's kernel estimator (2.12), we need to evaluate

$$\hat{\psi}(u_i) = \frac{\hat{f}^{(1)}(u_i)}{\hat{f}(u_i)} = \frac{1}{h^2}\frac{\sum_1^n (u_i - u)K(u_i - u)}{\sum_1^n K(u_i - u)}. \tag{3.6}$$

Further, if the Hjort and Glad kernel estimator with parametric start in (2.13) is used

$$\begin{aligned}\hat{\psi}(u_i) &= \frac{f^{(1)}(u_i,\eta)}{f(u_i,\eta)} + \frac{\hat{r}^{(1)}(u_i)}{\hat{r}(u_i)} \\ &= \psi(u_i,\eta) + \hat{\psi}_r(u_i),\end{aligned} \tag{3.7}$$

where $\psi(u_i,\eta)$ is the score of the parametric start $f(u_i,\eta)$, and $\hat{\psi}_r(u_i) = h^{-2}\sum_1^n (u_i - u)K(u_i - u)(f(u_i,\eta))^{-1}/\sum_1^n K(u_i - u)(f(u_i,\eta))^{-1}$ is the score of $\hat{r}(u_i)$.

3.2. Testing for heteroskedasticity

Let us consider the regression model

$$y_i = x_i'\beta + u_i, \quad u_i = \sigma_i v_i, \quad \sigma_i = \sigma(w_i,\alpha), \tag{3.8}$$

where v_i has an unknown density $f(v)$ and it is independent of both w_i and x_i, w_i can be the same or different from $x_i, \sigma(w_i,\alpha)$ is a known positive function of the vector w_i satisfying $\sigma(w_i,0) = \sigma$. We are interested in testing $\mathrm{H}_0{:}\,\alpha = 0$ against $\mathrm{H}_1{:}\,\alpha \neq 0$. The model (3.8) then satisfies (2.1) with $z_i=(y_i,x_i,w_i)$, $\theta=(\beta,\alpha)$, $v_1(z_i,\theta)=(y_i-x_i'\beta)/\sigma_i=v_i$, $v_2(x_i,w_i,\theta) = (x_i,w_i)$ and $J(z_i,\theta)$ is the Jacobian of the transformation from v_i to y_i. The efficient score function is $s(\theta) = \sum_1^n s(z_i,\theta,f)$ where

$$s(z_i,\theta,f) = -\begin{pmatrix} (\frac{x_i}{\sigma_i} - E(\frac{x_i}{\sigma_i}))\psi(v_i) \\ (1+\psi(v_i)v_i)(\frac{1}{\sigma_i}\frac{\partial\sigma_i}{\partial\alpha} - E(\frac{1}{\sigma_i}\frac{\partial\sigma_i}{\partial\alpha})) \end{pmatrix} = \begin{pmatrix} s_\beta(z_i,\theta,f) \\ s_\alpha(z_i,\theta,f)\end{pmatrix}, \tag{3.9}$$

$$s_\beta = -\left(\frac{x_i}{\sigma_i} - E\left(\frac{x_i}{\sigma_i}\right)\right)\psi(v_i)$$

[6] If the density is asymmetric and the regression model includes a constant, the efficient score for the constant is zero. The constant is not identifiable.

and

$$s_\alpha = -(1+\psi(v_i)v_i)\left(\frac{1}{\sigma_i}\frac{\partial\sigma_i}{\partial\alpha} - E\left(\frac{1}{\sigma_i}\frac{\partial\sigma_i}{\partial\alpha}\right)\right).$$

Thus locally most powerful test for heteroskedasticity is given by SPRS$=nR^2$ in the regression of unity on the score $s(z_i,\hat{\theta},\hat{f})$ evaluated under the null.

4. Monte Carlo simulation

We carry out a Monte Carlo simulation in order to assess, in small samples, the size and power of the SPRS tests discussed in the previous sections. We study the behavior of the tests on two instances: testing for omitted variables in the specification of the regression, and testing for heteroskedasticity.

In the first case, we design a regression model with two regressors

$$y_i = (x_{1i}-\bar{x}_1)\beta_1 + (x_{2i}-\bar{x}_2)\beta_2 + u_i, \tag{4.1}$$

where the hypothesis of interest is

$$\mathrm{H}_0\colon \beta_1 = 0, \qquad \mathrm{H}_1\colon \beta_1 \neq 0. \tag{4.2}$$

The regressors x_1 and x_2 (in deviations from their means) are drawn from a normal probability density function with zero mean and variance equal to one. The error term u_i is drawn from a Student-t with 5 degrees of freedom or from a Chi-squared with a 8 degrees of freedom.

In the second instance, the researcher wishes to test for heteroskedasticity in the error term of a regression model. The specification considered is

$$y_i = (x_{1i}-\bar{x}_1)\beta_1 + u_i, \quad u_i = \sigma_i v_i, \quad \sigma_i^2 = \sigma^2\exp(\alpha w_i), \tag{4.3}$$

where the hypothesis of interest is

$$\mathrm{H}_0\colon \alpha = 0, \qquad \mathrm{H}_1\colon \alpha \neq 0, \tag{4.4}$$

the error v_i is drawn from a Student-t with 5 degrees of freedom and the variable w_i is normally distributed with zero mean and variance equal to one.

In all experiments, we consider sample sizes of 25, 100, 250 and 500 observations and we perform 2,000 replications. We compare four tests: 1. PRS, parametric Rao's score test where the density is fully known. This is an ideal situation since in many instances the researcher deals with an unknown density, but we present the performance of this test as a benchmark. 2. QPRS, parametric Rao's score where the density is normal even though the true density may be something else. 3. SPRS$_k$, the SPRS test in (2.15) with kernel density estimator in (2.12). 4. SPRS$_{\mathrm{ps}}$, the SPRS test in (2.15) using the kernel estimator with a parametric start in (2.13).

In Table 1, we assess the size of the two semiparametric tests for omitted variables when the error term is distributed as a Student-t with 5 degrees of freedom. This implies a symmetric and leptokurtic error term with coefficient of kurtosis to equal to

Table 1
Size (%) of the semiparametric Rao score tests. Regression model with Student-t errors (t_5)[a]

$$y_i = (x_{1i} - \bar{x}_1)\beta_1 + (x_{2i} - \bar{x}_2)\beta_2 + u_i$$

$$\mathrm{H}_0\colon \beta_1 = 0 \quad \beta_1 \neq 0$$

	χ_1^2 c.v.	PRS	QPRS	SPRS_k	$\mathrm{SPRS}_{\mathrm{ps}}$
$n=25$	10%	12.15	12.45	11.80	11.20
	5%	5.90	5.30	5.50	5.60
	1%	1.05	0.50	0.90	0.90
	Mean = 1	1.08	1.15	1.04	1.09
	Variance = 2	2.11	1.90	1.85	1.90
$n=100$	10%	9.50	11.00	10.00	10.10
	5%	4.55	4.90	4.70	5.30
	1%	0.95	0.60	0.70	0.70
	Mean = 1	0.99	1.03	0.98	0.99
	Variance = 2	1.85	1.80	1.78	1.80
$n=250$	10%	9.80	10.50	10.10	11.10
	5%	5.05	5.20	4.60	5.30
	1%	1.00	0.90	0.90	0.90
	Mean = 1	1.01	1.02	0.99	1.01
	Variance = 2	1.99	2.06	1.86	1.98
$n=500$	10%	9.40	10.00	9.90	10.60
	5%	4.95	4.95	4.70	4.90
	1%	0.85	1.15	1.20	1.10
	Mean = 1	0.97	1.00	0.98	0.99
	Variance = 2	1.91	2.05	1.96	2.09

[a]PRS: Parametric RS test under the true Student-t errors. QPRS: Quasi Parametric RS test under misspecified normality. SPRS_k: Semiparametric RS test with kernel. $\mathrm{SPRS}_{\mathrm{ps}}$: Semiparametric RS test with kernel with a parametric start.

9. We compare them with the PRS test under the true density function (Student-t), and with a QPRS test under the assumption of normal errors. Under the null hypothesis, all tests are distributed asymptotically as a χ^2 with 1 degree of freedom. We assess the empirical size for the 10%, 5% and 1% significance levels of a χ_1^2 and report the empirical mean and variance of the tests in order to be compared with their theoretical values, which, under the null hypothesis, are 1 and 2, respectively. The general finding is that the four tests, PRS, QPRS, SPRS_k and $\mathrm{SPRS}_{\mathrm{ps}}$ have very good size behavior. The larger size distortions happen for small samples $n=25$ and 100, and for the QPRS. For samples $n=100$ and larger, the PRS has smaller size than the theoretical size. Both semiparametric tests have very similar size behavior with SPRS_k showing a smaller size than $\mathrm{SPRS}_{\mathrm{ps}}$. The latter exhibits a general tendency to have a slightly larger size than the theoretical one.

In Table 2, based on estimated critical values, we present the power performance of the PRS, QPRS, SPRS_k, and $\mathrm{SPRS}_{\mathrm{ps}}$ tests for omitted variables under the assumption that the errors are distributed as Student-t with 5 degrees of freedom. We consider two very local alternative hypothesis: $\beta_1 = 0.05$ and 0.1. The four tests are consistent since the power goes to one as the sample size increases. There are no extraordinarily big

Table 2
Power (%) of the semiparametric Rao score tests. Regression model with Student-t errors (t_5)

$$y_i = (x_{1i} - \bar{x}_1)\beta_1 + (x_{2i} - \bar{x}_2)\beta_2 + u_i$$

	Empirical size (%)	PRS	QPRS	SPRS_k	SPRS_{ps}
$H_0: \beta_1 = 0$	$H_1: \beta_1 = 0.05$				
$n = 25$	10	11.25	10.80	9.50	11.00
	5	5.15	6.05	5.70	6.20
	1	1.10	1.15	1.20	1.30
$n = 100$	10	14.30	13.25	11.60	13.10
	5	7.75	7.45	7.00	6.10
	1	2.00	1.60	1.60	1.80
$n = 250$	10	19.50	15.70	14.00	15.80
	5	10.60	9.90	8.80	9.30
	1	2.90	3.10	2.30	3.10
$n = 500$	10	26.55	23.05	20.50	23.00
	5	16.25	14.80	11.70	16.30
	1	6.85	4.25	2.90	4.70
$H_0: \beta_1 = 0$	$H_1: \beta_1 = 0.1$				
$n = 25$	10	13.60	12.75	11.00	13.20
	5	7.10	7.35	6.60	7.10
	1	1.55	1.45	1.80	2.70
$n = 100$	10	26.20	22.00	18.50	20.80
	5	17.15	14.60	10.40	13.00
	1	4.65	5.05	3.40	3.80
$n = 250$	10	40.30	34.50	29.00	33.70
	5	28.40	24.50	19.50	23.80
	1	11.60	10.75	7.30	9.60
$n = 500$	10	65.05	54.25	45.20	56.10
	5	50.85	42.45	34.00	44.90
	1	31.10	19.60	14.10	22.50

differences among the four tests in terms of power. PRS is the most powerful. Among the semiparametric tests, SPRS_{ps} is marginally more powerful than SPRS_k, and for large samples, $n = 500$, is slightly more powerful than QPRS.

In Tables 3 and 4, we present the results of these tests for the case in which the researcher wishes to test for heteroskedasticity under the assumption of Student-t distributed errors. In Table 3, we show the size performance. In general, the QPRS shows large size distortions and it is dominated by the semiparametric tests. For a small sample, $n = 25$, the four tests exhibit big size distortions, especially QPRS, SPRS_k, and SPRS_{ps}. In moderate samples, $n = 100$ and 250, PRS, QPRS and SPRS_{ps} have a larger size than the nominal size and SPRS_k has a smaller size than the nominal. Between the semiparametric tests, the SPRS_{ps} has better size than the SPRS_k, and it also dominates the QPRS test. For large samples, $n = 500$, the SPRS_k dominates both parametric and semiparametric tests, QPRS and SPRS_{sp}.

In Table 4, we present the power performance of these tests for the heteroskedasticity problem under two sets of alternative hypothesis. The main finding is that the QPRS

Table 3
Size (%) of the semiparametric Rao score tests. Regression model with heteroskedastic errors

$$y_i = (x_{1i} - \bar{x}_1)\beta_1 + u_i, \quad u_i = \sigma_i v_i, \quad \sigma_i^2 = \sigma^2 \exp(\alpha w_i)$$

v_i is Student-t (t_5) distributed error

$$H_0: \alpha = 0 \quad H_1: \alpha \neq 0$$

	χ_1^2 c.v.	PRS	QPRS	$SPRS_k$	$SPRS_{ps}$
$n=25$	10%	11.95	7.15	4.00	9.35
	5%	3.55	1.00	1.75	1.05
	1%	1.00	0.00	0.45	0.00
	Mean = 1	1.06	1.05	0.73	1.11
	Variance = 2	1.35	0.95	0.93	1.03
$n=100$	10%	11.25	12.35	7.65	11.10
	5%	5.75	6.00	3.20	5.30
	1%	1.15	1.00	0.10	0.90
	Mean = 1	1.06	1.16	0.90	1.12
	Variance = 2	2.04	1.92	1.25	1.86
$n=250$	10%	10.20	12.05	9.45	11.65
	5%	4.90	6.15	3.70	5.65
	1%	1.25	1.50	0.70	1.25
	Mean = 1	1.03	1.15	0.99	1.12
	Variance = 2	2.17	2.49	1.65	2.12
$n=500$	10%	10.70	12.20	10.10	12.40
	5%	6.35	6.80	5.70	7.05
	1%	1.65	2.05	1.30	2.15
	Mean = 1	1.07	1.18	1.03	1.21
	Variance = 2	2.43	2.95	2.27	2.76

looses power in the heteroskedasticity case with respect to the semiparametric test $SPRS_k$. Since for n=25 there are big size distortions, we focus our analysis in moderate and large samples. At the 10% level, the power of QPRS ranges from 7.60% to 26.75%, while that of $SPRS_k$ ranges from 15.90% to 41.50%, for the alternative hypothesis α=0.10. For the alternative α=0.25, the range of the power of QPRS goes from 21.45% to 84.70%, and the range of $SPRS_k$ goes from 39.20% to 90.85%. The semiparametric test $SPRS_k$ clearly dominates the parametric QPRS test. The semiparametric test $SPRS_{ps}$ has very similar power to the QPRS test.

In summary, the SPRS tests have very similar performance to the QPRS statistic in testing for omitted variables but they clearly dominate the QPRS test in testing for heteroskedasticity, since $SPRS_k$ is more powerful. An explanation for this behavior lies in the variance of the score function. For the case of omitted variables, we see that in (3.3) the variance of the score is proportional to $E(\psi_i^2)$. In the QPRS test, $E(\psi_i^2) = E(u_i^2) = \sigma^2$. In our Monte Carlo experiment, the true density function is a Student-t with 5 degrees of freedom, so that $\sigma^2 = 1.67$. The PRS test uses the true score, and consequently for a Student t, $E(\psi_i^2) = (\gamma+1)/(\gamma+3)$ where γ is the degrees of freedom. If $\gamma = 5$, $E(\psi_i^2) = 0.75$. Comparing the variance of the score in the QPRS and in the PRS tests, we observe that the former is 2.2 times larger than the latter.

Table 4
Power (%) of the semiparametric Rao score tests. Regression model with heteroskedastic errors

$$y_i = (x_{1i} - \bar{x}_1)\beta_1 + u_i, \quad u_i = \sigma_i v_i, \quad \sigma_i^2 = \sigma^2 \exp(\alpha w_i)$$

v_i is Student-t (t_5) distributed error

	Empirical size (%)	PRS	QPRS	SPRS_k	SPRS_{ps}
$H_0: \alpha = 0$ $H_1: \alpha = 0.1$					
$n = 25$	10	6.25	5.85	16.95	6.05
	5	3.20	3.40	9.80	3.35
	1	0.60	0.75	1.85	0.65
$n = 100$	10	11.55	7.60	15.90	6.55
	5	4.60	2.30	8.50	2.35
	1	0.30	0.20	1.90	0.20
$n = 250$	10	26.10	13.30	23.80	11.50
	5	16.65	4.65	15.15	4.85
	1	1.85	0.20	2.55	0.30
$n = 500$	10	51.50	26.75	41.50	22.75
	5	32.80	10.25	25.00	8.75
	1	8.20	0.70	6.15	0.75
$H_0: \alpha = 0$ $H_1: \alpha = 0.25$					
$n = 25$	10	7.50	6.40	27.05	5.05
	5	2.85	3.85	18.35	2.50
	1	0.50	1.10	4.75	0.75
$n = 100$	10	45.30	21.45	39.20	17.65
	5	21.85	8.55	27.30	6.90
	1	3.40	1.10	9.60	0.55
$n = 250$	10	87.55	54.90	67.85	44.80
	5	76.30	29.65	56.45	24.40
	1	32.75	4.10	25.90	3.65
$n = 500$	10	99.55	84.70	90.85	73.25
	5	98.20	64.30	83.05	45.25
	1	86.35	20.40	55.20	12.60

In the heteroskedasticity experiment, according to (3.9), the variance of the score is proportional to $E(1 + \psi(v_i)v_i)^2$. In the QPRS test, $E(1 + \psi(v_i)v_i)^2 = k - 1$, where k is the coefficient of kurtosis of the underlying density function. In our Monte Carlo experiment, the true density function is a standardized Student-t with 5 degrees of freedom, for which $k = 9$ and $E(1 + \psi(v_i)v_i)^2 = 8$. This QPRS test due to Koenker (1981) is the robust version of Breusch and Pagan (1979) test where the normal density ($k = 3$) is used. The PRS test uses the true score function, where, for a standardized Student-t, $E(1+\psi(v_i)v_i)^2 = 2\gamma/(\gamma+3)$, and for $\gamma = 5$, $E(1+\psi(v_i)v_i)^2 = 1.25$. Comparing the variances of the score in QPRS and PRS tests, we observe that the former is 6.4 time larger than the latter. The variances of the scores of semiparametric tests lie between those of the PRS and QPRS tests. In the heteroskedasticity case, the QPRS test becomes very inefficient in comparison with that of the omitted variables case, leaving enough potential gain for the SPRS test. For a theoretical comparison of the asymptotic

Table 5
Size (%) of the semiparametric Rao score tests. Regression model with chi-squared errors (χ^2_8)

$$y_i = (x_{1i} - \bar{x}_1)\beta_1 + (x_{2i} - \bar{x}_2)\beta_2 + u_i$$

$$H_0: \beta_1 = 0 \quad \beta_1 \neq 0$$

	χ^2_1 c.v.	PRS	QPRS	SPRS$_k$
$n = 25$	10%	8.70	16.00	11.95
	5%	4.75	9.35	6.05
	1%	1.50	2.57	1.10
	Mean = 1	0.97	1.36	1.07
	Variance = 2	2.51	3.25	2.11
$n = 100$	10%	8.75	10.75	10.75
	5%	4.90	6.25	4.95
	1%	1.50	1.35	0.70
	Mean = 1	1.01	1.07	1.00
	Variance = 2	2.78	2.42	1.85
$n = 250$	10%	9.85	10.15	9.15
	5%	5.05	4.90	4.75
	1%	1.50	1.25	0.95
	Mean = 1	1.01	1.01	0.99
	Variance = 2	2.97	2.02	1.88
$n = 500$	10%	10.50	10.40	10.55
	5%	6.00	5.70	6.05
	1%	1.75	1.65	0.85
	Mean = 1	1.06	1.05	1.04
	Variance = 2	2.79	2.43	2.10

variances of the maximum likelihood, quasi-maximum likelihood and semiparametric estimators, see González-Rivera and Drost (1999).

Finally, in Tables 5 and 6, we present the results of PRS, QPRS and SPRS for the case of testing for omitted variables under the assumption that the errors are distributed as χ^2 with 8 degrees of freedom. This implies an asymmetric and leptokurtic error term with coefficient of skewness equal to 1 and a coefficient of kurtosis equal to 4.5. In Table 5, we show the size of the tests. For small samples, $n = 25$, there is a big size distortion in the QPRS test. For moderate samples, $n = 100$, the SPRS$_k$ has a better size than the QPRS. For moderate to large samples, $n = 250$ and 500, there is a similar performance in QPRS and SPRS$_k$. In Tables 6, we report the power performance for two sets of alternative hypothesis, $\beta_1 = 0.1$ and 0.4. Since for $n = 25$, there are big size distortions, we focus in moderate and large samples. For the very local alternative $\beta_1 = 0.1$, there is a very similar performance in QPRS and SPRS$_k$ where the former dominates the latter. For the alternative $\beta_1 = 0.4$, we observe that there are some small power gains in the SPRS$_k$ with respect to the QPRS, in the moderate to large samples, $n = 100$, 250 and 500.

In summary, these Monte Carlo simulations show that there is a positive pay-off to the use of the SPRS tests. For the problem of misspecified regression, the semiparametric tests perform as well as the QPRS test. For the problem of heteroskedasticity, the

Table 6
Power (%) of the semiparametric Rao score tests regression model with chi-squared errors (χ^2_8)

$$y_i = (x_{1i} - \bar{x}_1)\beta_1 + (x_{2i} - \bar{x}_2)\beta_2 + u_i$$

	Empirical size (%)	PRS	QPRS	SPRS_k
$H_0\colon \beta_1 = 0$ $H_1\colon \beta_1 = 0.1$				
$n = 25$	10	7.05	9.72	5.30
	5	2.45	5.09	2.85
	1	0.35	1.72	0.65
$n = 100$	10	10.10	9.85	6.80
	5	3.20	4.30	4.20
	1	0.05	0.08	1.20
$n = 250$	10	11.85	13.10	8.90
	5	5.30	7.10	4.75
	1	0.25	1.00	1.40
$n = 500$	10	21.65	16.75	11.40
	5	9.85	8.65	6.40
	1	0.95	1.20	2.05
$H_0\colon \beta_1 = 0$ $H_1\colon \beta_1 = 0.4$				
$n = 25$	10	8.50	14.00	11.40
	5	1.65	7.21	6.60
	1	0.00	2.57	1.65
$n = 100$	10	41.90	26.50	31.45
	5	20.90	14.00	22.30
	1	0.25	4.20	9.15
$n = 250$	10	77.45	61.80	58.00
	5	59.55	47.80	43.05
	1	7.80	20.75	24.60
$n = 500$	10	93.40	87.80	85.50
	5	87.80	79.50	75.20
	1	54.60	45.85	55.75

semiparametric tests outperformed the QPRS test. Evaluating the overall performance of the two semiparametric tests considered we recommend the use of SPRS_k become of its superior power properties.

Acknowledgements

We are grateful to A. Bera, A. Pagan, J. Powell and four referees for their useful comments and discussions. We are especially thankful to one of the referees for his detailed and constructive comments and suggestions. The comments from the participants at McGill University, University of Victoria, Bilkent University, American Statistical Association meeting, and the Bernoulli Society meeting also were very helpful. The authors gratefully acknowledge research support from the Academic Senate, UCR.

Appendix A.

We show that the SPRS has a central χ_q^2 distribution under the null. The proof, using the assumptions 1–14 of Ai (1997) and the assumption that $g(\theta)$ is twice continuously differentiable with $G(\theta) = \partial g(\theta)/\partial\theta'$ nonsingular, follows from the results of Newey and McFadden (1994) and Ai (1997).

First we note from Lemmas $A.1$ and $A.2$ of Ai (1997) that, uniformly over θ,

$$\frac{1}{n}\hat{s}(\theta) = E[s(z,\theta,f)] + \mathrm{o_p}(1), \tag{A.1}$$

$$\frac{1}{n}\frac{\partial}{\partial\theta'}\hat{s}(\theta) = E\left[\frac{\partial}{\partial\theta'}s(z,\theta,f)\right] + \mathrm{o_p}(1) = -\Sigma + \mathrm{o_p}(1) \tag{A.2}$$

or

$$\frac{1}{n}\frac{\partial}{\partial\theta'}(\hat{s}(\theta) - s(\theta)) = \mathrm{o_p}(1),$$

$$\frac{1}{\sqrt{n}}(\hat{s}(\theta) - s(\theta)) = \mathrm{o_p}(1), \tag{A.3}$$

$$\frac{1}{n}\sum_1^n s(z_i,\theta,\hat{f})s(z_i,\theta,\hat{f})' = \Sigma + \mathrm{o_p}(1), \tag{A.4}$$

where $\hat{s}(\theta) = \sum_1^n s(z_i,\theta,\hat{f})$. Under assumption 8 of Ai (1997) $Es(z,,\theta,f) = 0$ if and only if $\theta=\theta_0$, the result (A.1) gives $\hat{s}(\theta)/n \xrightarrow{p} 0$. Further from (A.2) $n^{-1}\partial\hat{s}(\theta)/\partial\theta \xrightarrow{p} -\Sigma$. Thus $\hat{\theta}$ is consistent, $\hat{\theta} = \theta + \mathrm{o_p}(1)$ under the null.

Now substituting (A.2) and (A.3) in (2.17) and using (2.11)

$$\sqrt{n}\hat{\theta} = \sqrt{n}\tilde{\theta} + \mathrm{o_p}(1) \quad \text{and} \quad \sqrt{n}(\hat{\lambda} - \tilde{\lambda}) = \mathrm{o_p}(1). \tag{A.5}$$

Further, using a Taylor expansion around f and $\tilde{\theta}$

$$\frac{1}{\sqrt{n}}\hat{s}(\hat{\theta}) = \frac{1}{\sqrt{n}}s(\tilde{\theta}) + \mathrm{o_p}(1). \tag{A.6}$$

These results show that

$$\mathrm{SPRS} = \mathrm{SPRS}^* + \mathrm{o_p}(1). \tag{A.7}$$

Thus, under the null, the asymptotic distribution of SPRS is a central χ_q^2. The results follows immediately from the Newey and McFadden (1994) result that the SPRS* is distributed as a central χ_q^2.

Finally we show that the one-step estimator $\hat{\theta}^*$ is asymptotically equivalent to $\hat{\theta}$ under the null. This follows by first noting that $\partial\hat{s}(\theta^*)/n\partial\theta' = -\Sigma + \mathrm{o_p}(1)$ and $\partial g(\theta^*)/\partial\theta' = G + \mathrm{o_p}(1)$. Further, by Taylor expansion, $\sqrt{n}g(\theta^*) = \sqrt{n}g(\theta) + G\sqrt{n}(\theta^* - \theta) + \mathrm{o_p}(1)$ and $(1/\sqrt{n})\hat{s}(\theta^*) = (1/\sqrt{n})\hat{s}(\theta) - \Sigma\sqrt{n}(\theta^* - \theta) + \mathrm{o_p}(1)$. Substituting these into (2.17) and then comparing with (2.11) prove that $\sqrt{n}(\hat{\theta}^* - \tilde{\theta}) = \mathrm{o_p}(1)$ and $\sqrt{n}(\hat{\lambda}^* - \tilde{\lambda}) = \mathrm{o_p}(1)$. Thus, $\sqrt{n}(\hat{\theta}^* - \hat{\theta}) = \sqrt{n}(\hat{\theta}^* - \tilde{\theta}) + \sqrt{n}(\tilde{\theta} - \hat{\theta}) = \mathrm{o_p}(1)$.

References

Ai, C., 1997. A semiparametric maximum likelihood estimator. Econometrica 97, 933–963.

Begun, J.M., Hall, W., Huang, W.M., Wellner, J.A., 1983. Information and asymptotic sufficiency in parametric-non parametric models. Ann. Statist. 11, 432–452.

Bera, A.K., Bilias, Y., 1997. Rao's score, Neyman's C(α) and Silvey's LM tests: an essay on historical developments and some new results. Working paper, University of Illinois, Urbana-Champaign.

Bera, A.K., Ullah, A., 1991. Rao's score test in econometrics. J. Quant. Econom. 7, 189–220.

Bickel, P.J., 1978. Using residuals robustly; test for heteroskedasticity and nonlinearity. Ann. Statist. 6, 266–291.

Bickel, P.J., 1982. On adaptive estimation. Ann. Statist. 10, 647–671.

Bickel, P.J., Klassen, C.A.J., Ritov, Y., Wellner, J.A., 1993. Efficient and Adaptive Estimation for Semiparametric Models. The John Hopkins University Press, Baltimore, Maryland.

Breusch, T.S., Pagan, A.R., 1979. A simple test for heteroskedasticity and random coefficient variation. Econometrica 47, 1287–1294.

Godfrey, L.G., 1988. Misspecification Tests in Econometrics, The Lagrange Multiplier Principle and other Approaches. Cambridge University Press, New York.

González-Rivera, G., Drost, F.C., 1999. Efficiency comparisons of maximum likelihood-based estimators in GARCH models. Journal of Econometrics 93, 93–111.

Hjort, N.L., Glad, I.K., 1995. Non parametric density estimation with a parametric start. Ann. Statist. 23, 882–904.

Ichimura, H., Lee, L.F., 1991. Semiparametric least square estimation of multiple models. In: Barnett, W., Powell, J., Tauchen, G. (Eds.), Non Parametric and Semiparametric Methods in Econometrics and Statistics. Cambridge University Press, Cambridge.

Koenker, R., 1981. A note on studentizing a test for heteroscedasticity. J. Econometrics 17, 107–112.

Kreiss, J.P., 1990. Testing linear hypotheses in autoregressions. Ann. Statist. 18, 1470–1482.

Marron, J.S., 1988. Automatic smoothing parameter selection: a survey. In: Ullah, A. (Ed.), Semiparametric and Nonparametric Econometrics. Physica-Verlag, Heidelberg.

Newey, W.K., 1994. The asymptotic variance of semiparametric estimators. Econometrica 62, 1349–1382.

Newey, W.K., McFadden, D., 1994. Large sample estimation and hypothesis testing. In: Engle, R.F., McFadden, D.L. (Eds.), Handbook of Econometrics, Vol. IV. Elsevier Science, B.V., Amsterdam (Chapter 36).

Pagan, A., Ullah, A., 1996. Nonparametric econometrics. Manuscript, University of California, Riverside, Department of Economics.

Robinson, P.M., 1988. Root-N-consistent semiparametric regression. Econometrica 56, 931–954.

Rosenblatt, M., 1956. Remarks on some nonparametric estimates of density function. Ann. Math. Statist. 27, 832–837.

Silverman, B.W., 1986. Density Estimation for Statistics and Data Analysis. Chapman & Hall, London.

Silvapulle, M.J., Silvapulle, P., Basawa, I.V., 1997. On adaptive tests. Conference Proceedings of the Econometric Society Australasian Meeting, Vol. 2, pp. 480–499.

Stone, C.J., 1975. Adaptive maximum likelihood of a location parameter. Ann. Statist. 3, 267–284.