Ukrainian Mathematical Journal, Vol. 56, No. 4, 2004

A GOODNESS-OF-FIT TEST FOR A POLYNOMIAL ERRORS-IN-VARIABLES MODEL

C.-L. Cheng [1] **and A. G. Kukush** [2]

UDC 519.21

Polynomial regression models with errors in variables are considered. A goodness-of-fit test is constructed, which is based on an adjusted least-squares estimator and modifies the test introduced by Zhu et al. for a linear structural model with normal distributions. In the present paper, the distributions of errors are not necessarily normal. The proposed test is based on residuals, and it is asymptotically chi-squared under null hypothesis. We discuss the power of the test and the choice of an exponent in the exponential weight function involved in test statistics.

1. Introduction

Cheng and Schneeweiss [1] developed an adjusted least-squares (ALS) estimator for the parameters of a polynomial functional regression model with errors in variables. The estimator is consistent and asymptotically normal and can be viewed as resulting from the principle of corrected unbiased estimating equations (see, e.g., [2, Chap. 6]). In [3], a small sample modification of the ALS estimator was constructed, which shows good results in small sample and is asymptotically equivalent to the ALS estimator. For a further discussion of related models, see [1, 4].

Errors-in-variables (EIV) models are widely used in practical applications. Therefore, it is reasonable to develop appropriate goodness-of-fit tests. However, the literature in the EIV context mostly deals with estimation rather than testing. In [5], a goodness-of-fit test based on residuals was presented for a linear structural EIV model, where the distribution of the latent variable and the error distributions were normal. The normality assumption was crucial for correcting the bias of a test of the score type.

In the present paper, we modify that goodness-of-fit test for polynomial functional relations. We assume that the measurement errors possess finite exponential moments and use an exponential weight function in the test statistics. The bias correction of the test is now performed on the basis of the exponential moments of errors, which are supposed to be known. The test relies on the ALS estimator and its small sample modification.

In the present paper, we use a standard notation: $\mathbf{E}\varepsilon$ and $\operatorname{var}(\varepsilon)$ denote the expectation and the variance of a random variable ε, $\operatorname{cov}(\xi)$ is the variance-covariance matrix of a random vector ξ, $O_p(1)$ denotes a sequence of stochastically bounded random variables, and $o_p(1)$ is a sequence of random variables that converges to 0 in probability.

In Sec. 2, we describe the model and the ALS estimator. In Sec. 3, we present the goodness-of-fit test and show that it is asymptotically chi-squared under the null hypothesis. We introduce local alternatives and investigate the power of the test in Sec. 4. There we also discuss an optimal choice of the exponent in the weight function of test statistics. Section 5 distinguishes two important particular cases where the assumptions are simplified: (a) linear functional model and (b) polynomial structural model, where the latent variable is random and has an unknown distribution. Section 6 is concluding, and the proofs of the results are presented in Appendix.

[1] Academia Sinica, Taipei, Republic of China.

[2] Shevchenko Kiev University, Kiev.

Published in Ukrains'kyi Matematychnyi Zhurnal, Vol. 56, No. 4, pp. 527–543, April, 2004. Original article submitted October 3, 2002.

0041–5995/04/5604–0641

2. Model and Estimator

We discuss the polynomial functional relation

$$y_i = \zeta_i'\beta + \varepsilon_i, \tag{1}$$

$$x_i = \xi_i + \delta_i, \tag{2}$$

where $\zeta_i' := \left(1, \xi_i, \xi_i^2, \ldots, \xi_i^k\right)$, $k \ge 1$, $\beta := (\beta_0, \beta_1, \ldots, \beta_k)'$, and ε_i, $i = 1, \ldots, n$, are the n sample values of a latent nonstochastic variable ξ. The sequences δ_i, $i = 1, \ldots, n$, and ε_i, $i = 1, \ldots, n$, are two IID sequences of random errors, independent of each other, with expectation 0. We assume that $\sigma_\varepsilon^2 := \mathbf{E}\varepsilon^2 < \infty$, and admit the particular case $\sigma_\varepsilon^2 = 0$, which means that the response variable y can be observed without error. However, we assume that $\sigma_\delta^2 := \mathbf{E}\delta^2 > 0$. The variance σ_ε^2 of ε need not be known, but it is assumed that all moments $\mathbf{E}(\delta^l)$, $l = 1, \ldots, 2k$, are known; moreover, we need some exponential moments of δ.

For the observable x, let $t_r(x)$ be a polynomial of degree r such that $\mathbf{E}\, t_r(x) = \mathbf{E}\, t_r(\xi + \delta) = \xi^r$, $r = 0, 1, \ldots, 2k$. The polynomial $t_r(x)$ can be expressed via the moments $\mathbf{E}\delta^l$, $l = 1, \ldots, r$ (see [1]). Denote $t = t(x) := \left(t_0(x), t_1(x), \ldots, t_k(x)\right)'$ and let $H = H(x)$ be a $(k+1) \times (k+1)$ matrix the (p, q) element of which is t_{p+q}, $p, q = 0, \ldots, k$. The ALS estimator $\hat\beta$ of β satisfies the equation

$$\overline{H}\hat\beta = \overline{ty}. \tag{3}$$

Hereafter, the overbars denote averages, i.e.,

$$\overline{ty} := \frac{1}{n}\sum_{i=1}^{n} t(x_i)\, y_i,$$

etc. For an arbitrary function f, we denote $\mathcal{M}(f(\xi)) := \lim_{n\to\infty} \overline{f(\varepsilon)}$, provided that the limit exists and is finite.

Lemma 1 [1]. *Assume that the following conditions are satisfied:*

(i) $\mathbf{E}\delta^{4k} < \infty$;

(ii) the limit $\mathcal{M}(\xi^r)$ *exists for* $r = 1, \ldots, 4k$;

(iii) the matrix $S := \mathcal{M}[\zeta(\xi)\zeta'(\xi)]$ *is nonsingular.*

Then $\overline{H}$ *is nonsingular with probability tending to* 1 *and* $\hat\beta \xrightarrow{P} \beta$ *as* $n \to \infty$. *Moreover, the ALS estimator*

$$\hat{\sigma}_\varepsilon^2 := \overline{y^2} - \left(\overline{ty}\right)'\hat{\beta} \tag{4}$$

converges in probability to σ_ε^2 *as* $n \to \infty$.

3. Construction of the Test and Bias Correction

For the response variable y and the corresponding latent variable ξ, we consider the following hypotheses for fixed $k \geq 1$:

$$H_0: \text{ for some } \beta_0, \dots, \beta_k, \quad \mathbf{E}(y - \beta_0 - \beta_1\xi - \dots - \beta_k\xi^k) = 0 \tag{5}$$

versus

$$H_1: \text{ for all } \beta_0, \dots, \beta_k, \quad \mathbf{E}(y - \beta_0 - \beta_1\xi - \dots - \beta_k\xi^k) \text{ is not identically equal to } 0. \tag{6}$$

If we want to use the residuals in a test statistic for the hypothesis H_0 based on observed y's and x's, then we have to consider the expectation of the residual with x instead of ξ in (5). However, as discussed in [5], one has

$$\mathbf{E}(y - \beta_0 - \beta_1 x - \dots - \beta_k x^k) \neq 0$$

even if relation (5) is true. Therefore, a bias correction is needed.

We perform a correction as follows: Let $w(\cdot)$ be a weight function. Then the hypothesis H_0 yields

$$\mathbf{E}[(y - \zeta'\beta)w(\xi)] = 0. \tag{7}$$

We want to construct polynomials $s_0(x), s_1(x), \dots, s_k(x)$ such that, under the hypothesis H_0, the following relation is true:

$$\mathbf{E}[(y - s(x)'\beta)w(x)] = 0, \tag{8}$$

where $s(x) = s = (s_0(x), \dots, s_k(x))'$. It is possible to satisfy (8) if one chooses $w(x) = e^{\lambda x}$ with fixed $\lambda \neq 0$, provided that the corresponding exponential moments of δ exist. Assume that the following conditions is satisfied:

(iv) $\mathbf{E}[(1 + |\delta|^k)e^{\lambda\delta}] < \infty$.

Denote $\mu_r := \mathbf{E}(\delta^r e^{\lambda\delta})$. For the weight function chosen, relation (8) holds if, for every ξ, one has

$$\xi^r \mathbf{E} e^{\lambda\delta} = \mathbf{E}(s_r(\xi + \delta)e^{\lambda\delta}), \qquad r = 0, \dots, k. \tag{9}$$

We have $s_0(x) = 1$ and $s_1(x) = x - \mu_1/\mu_0$. We seek $s_r(x)$ in the form

$$s_r(x) \;=\; \sum_{j=0}^{r} b_{rj}\, x^j .$$

To satisfy (9), we have the relations

$$b_{rr} \;=\; 1, \qquad b_{rp} \;=\; -\frac{1}{\mu_0}\sum_{j=p+1}^{r}\binom{j}{p}\mu_{j-p}\, b_{rj}, \qquad p \;=\; r-1, r-2, \dots, 0, \tag{10}$$

which enables us to derive all coefficients in succession. In other words, the vector $b_r := (b_{r0},\, b_{r1}, \dots, b_{rr})'$ satisfies the set of equations

$$A b_r \;=\; e_r . \tag{11}$$

Here, $e_r = (0, 0, \dots, 1)' \in \mathbf{R}^{(r+1)\times 1}$ and $A \in \mathbf{R}^{(r+1)\times(r+1)}$ is an upper triangular matrix with the entries

$$a_{pj} \;=\; \binom{j}{p}\mu_{j-p}, \qquad 0 \;\le\; p \;\le\; j \;\le\; r. \tag{12}$$

Then $b_r = A^{-1} e_r$. Thus, relation (8) holds with

$$s_r(x) \;=\; \sum_{j=0}^{r-1} b_{rj}\, x^j \;+\; x^r, \tag{13}$$

where $b_{rj} = b_{rj}(\mu_1/\mu_0, \dots, \mu_{r-j}/\mu_0)$ are polynomial functions of the ratios.

With polynomials (13), we now consider the following statistic of the score type:

$$T_{n0} \;:=\; \frac{1}{n}\sum_{i=1}^{n}\Big(y_i - s(x_i)'\hat{\beta}\Big)e^{\lambda x_i} \;=\; \overline{\Big(y - s'\hat{\beta}\Big)e^{\lambda x}} . \tag{14}$$

Recall that the overbar denotes an average and $\hat{\beta}$ is the ALS estimator given in (3). We need further assumptions to derive an asymptotic expansion of $\sqrt{n}\,T_{n0}$:

(v) $\mathbf{E}[(1+\delta^{2k})e^{2\lambda\delta}] \;<\; \infty$ [this condition is stronger than (iv)];

(vi) $\mathcal{M}(\xi^r e^{\lambda\xi})$ exists for $r \;=\; 0, 1, \dots, 2k$;

(vii) $\overline{\xi^r} \;-\; \mathcal{M}(\xi^r) \;=\; o\big(n^{-1/4}\big)$ as $n \to \infty$ for $r \;=\; 1, \dots, 2k$.

Remark 1. If ξ_i are IID random variables with $\mathbf{E}\,|\,\xi\,|^{8k+\alpha} < \infty$, where $\alpha > 0$ fixed, then relation (vii) holds a.s. Indeed, by the law of large numbers, we have $\mathcal{M}(\xi^r) = \mathbf{E}\,\xi^r$ a.s., and, by virtue of the Rosenthal moment inequality [6],

$$\mathbf{E}\left|\frac{1}{n}\sum_{i=1}^{n}\left(\xi_i^r - \mathbf{E}\xi^r\right)\right|^{4+\alpha/r} \le \frac{\mathrm{const}\,\mathbf{E}|\xi|^{4r+\alpha}}{n^{2+\alpha/2r}}.$$

Therefore,

$$\mathbf{E}\left|n^{1/4}\left(\overline{\xi^r} - \mathbf{E}\xi^r\right)\right|^{4+\alpha/r} \le \frac{1}{n^{1+\alpha/4r}},$$

and

$$\sum_{n=1}^{\infty}\frac{1}{n^{1+\alpha/4r}} < \infty,$$

and relation (vii) holds a.s. by virtue of the Chebyshev inequality and the Borel–Cantelli lemma.

Remark 1 shows that condition (vii) is realistic; it is satisfied a.s. for a structural polynomial EIVM if ξ has finite higher moments.

Lemma 2. *Assume (i) to (iii) and (v) to (vii). Then*

$$\sqrt{n}T_{n0} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\varepsilon_i\left(e^{\lambda x_i} - t(x_i)'f\right) + \beta'\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_i + o_P(1), \tag{15}$$

where η_i *are independent random vectors with expectation* 0 *and*

$$\eta_i = (\zeta_i - s(x_i))e^{\lambda x_i} + (H_i - \zeta_i t(x_i)')f, \tag{16}$$

$$f := S^{-1}\mathcal{M}(\zeta(\xi)e^{\lambda\xi})\mu_0,$$

where S *comes from (iii) and* $\mu_0 = \mathbf{E}e^{\lambda\delta}$.

We now introduce some more assumptions in order to apply the central limit theorem in the Lyapunov form to the sum of η_i:

(viii) for fixed $\alpha > 0$, $\mathbf{E}\left[\left(1+|\delta|^{2k+\alpha}\right)e^{(2+\alpha)\lambda\delta}\right] < \infty$ and $\mathbf{E}|\varepsilon|^{2+\alpha} < \infty$;

(ix) $\mathcal{M}(\xi^r e^{\lambda\xi})$ exists for $r = 0, 1, \dots, 3k$, and $\mathcal{M}(\xi^r e^{2\lambda\xi})$ exists for $r = 0, 1, \dots, 2k$;

(x) for fixed $\alpha > 0$, $\overline{|\xi|^{4k+\alpha}} + \overline{e^{\lambda(2+\alpha)\xi}} + \overline{|\xi|^{2k+\alpha}e^{\lambda(2+\alpha)\xi}} \le \mathrm{const}.$

Condition (viii) absorbs conditions (iv) and (v), and condition (ix) absorbs condition (vi). Condition (x) means that the higher empirical moments of ξ are bounded.

Lemma 3. *Assume (i) to (iii) and (vii) to (x). Then* $\sqrt{n}\,T_{n0} \xrightarrow{d} N\left(0, \sigma_T^2\right)$, *where*

$$\sigma_T^2 \;:=\; \sigma_\varepsilon^2 \mathcal{M}\left[\mathbf{E}(e^{\lambda x} - t'f)^2\right] + [\beta', f' \otimes \beta']\mathcal{M}\left(\operatorname{cov}\begin{pmatrix}(\zeta - s)e^{\lambda x}\\ \operatorname{vec}(H) - \operatorname{vec}(\zeta t')\end{pmatrix}\right)\begin{bmatrix}\beta\\ f \otimes \beta\end{bmatrix},$$

$f := S^{-1}\mathcal{M}(\zeta(\xi)e^{\lambda\xi})\mu_0$, $\otimes$ *is the Kronecker product,*

$$\mathcal{M}(\operatorname{cov}(Z(\xi, \delta))) \;:=\; \lim_{n\to\infty} \overline{\operatorname{cov}(Z(\xi_i, \delta))},$$

and $Z(\xi, \delta)$ *is a vector function of* ξ *and* δ.

Under the conditions of Lemma 3, the approximation of σ_T^2 is given by

$$A_n^2 \;:=\; \hat{\sigma}_\varepsilon^2 \overline{(e^{\lambda x} - t'\hat{f})^2} + \left[\hat{\beta}', \hat{f}' \otimes \hat{\beta}'\right]\widehat{\operatorname{cov}}\begin{pmatrix}(\zeta - s)e^{\lambda x}\\ \operatorname{vec}(H) - \operatorname{vec}(\zeta t')\end{pmatrix}\begin{bmatrix}\hat{\beta}\\ \hat{f} \otimes \hat{\beta}\end{bmatrix}. \tag{17}$$

Here, $\hat{f}$ and $\widehat{\operatorname{cov}}$ are approximations described below.

A. $\hat{f} = \overline{H}^{-1}\overline{se^{\lambda x}}$ because $\overline{H} \xrightarrow{P} S$, and

$$p\lim_{n\to\infty} \overline{s(x)e^{\lambda x}} = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \mathbf{E}\left(s(\xi_i + \delta)e^{\lambda(\xi_i + \delta)}\right) = \mu_0 \lim_{n\to\infty} \overline{\zeta(\xi)e^{\lambda\xi}} = \mu_0\mathcal{M}(\zeta(\xi)e^{\lambda\xi}).$$

B. We now have

$$\operatorname{cov}\begin{pmatrix}(\zeta - s)e^{\lambda x}\\ \operatorname{vec}(H) - \operatorname{vec}(\zeta t')\end{pmatrix} = \begin{pmatrix}\Sigma_{11}(\xi) & \Sigma_{12}(\xi)\\ \Sigma_{12}'(\xi) & \Sigma_{22}(\xi)\end{pmatrix}$$

Let us describe the approximations $\hat{\Sigma}_{ij}$ to $\mathcal{M}(\Sigma_{ij}(\xi))$. We get

$$\Sigma_{11} = \mathbf{E}(s(x)s(x)'e^{2\lambda x}) + \zeta\zeta' e^{2\lambda\xi}\mathbf{E}e^{2\lambda\delta} - \zeta e^{2\lambda\xi} - \mathbf{E}(s(x)'e^{2\lambda\delta})$$

$$- \mathbf{E}(s(x)e^{2\lambda\delta})e^{2\lambda\xi}\zeta' := U_1 + U_2 - U_3 - U_3'.$$

Then an approximation to $\mathcal{M}(U_1(\xi))$ is given by $\hat{U}_1 = \overline{s(x)s(x)'e^{2\lambda x}}$. We denote by $\tilde{s}_r(x)$, $r = 0, 1, \ldots, 2k$, the polynomials given in (13), but constructed for the exponent $\tilde{\lambda} := 2\lambda$. Thus,

$$\xi^r\mathbf{E}e^{2\lambda\delta} = \mathbf{E}\left(\tilde{s}_r(\xi + \delta)e^{2\lambda\delta}\right), \qquad r = 0, 1, \ldots, 2k.$$

Then $\hat{U}_2 = \overline{\left(\tilde{s}_{i+j}(x)e^{2\lambda x}\right)_{i,j=0}^{k}}$. The entries of U_3 can now be transformed to the sums of the values $\xi^r \mathbf{E} e^{2\lambda x}$, and $p\lim_{n\to\infty} \overline{\tilde{s}_r(x)e^{2\lambda x}} = \mathcal{M}(\xi^r e^{2\lambda x})$, $r = 0, \dots, 2k$. In U_3, we replace the summands $\xi^r \mathbf{E} e^{2\lambda x}$ by $\overline{\tilde{s}_r(x)e^{2\lambda x}}$, thereby obtaining $\hat{U}_3$. Finally, $\hat{\Sigma}_{11} = \hat{U}_1 + \hat{U}_2 - \hat{U}_3 - \hat{U}_3$. Next,

$$\Sigma_{12}' = \mathbf{E}\left(\operatorname{vec}(H)e^{\lambda x}\zeta'\right) - \mathbf{E}\left(\operatorname{vec}(H)e^{\lambda x}s(x)'\right) - \mathbf{E}\left(\operatorname{vec}(\zeta t')e^{\lambda x}\zeta'\right) + \mathbf{E}\left(\operatorname{vec}\left(\zeta t(x)'\right)e^{\lambda x}s(x)'\right)$$

$$:= V_1 - V_2 - V_3 + V_4.$$

We now have $\hat{V}_2 = \overline{\operatorname{vec}(H)e^{\lambda x}\, s(x)'}$. The entries of V_1, V_3, and V_4 can be transformed to the sums of the values $\xi^r \mathbf{E} e^{\lambda x}$. However, $p\lim_{n\to\infty} \overline{s_r(x)e^{\lambda x}} = \mathcal{M}(\xi^r \mathbf{E} e^{\lambda x})$, and we construct the further approximation by replacing $\xi^r \mathbf{E} e^{\lambda x}$ by $\overline{s_r(x)e^{\lambda x}}$. Finally, $\hat{\Sigma}_{12}' = \hat{V}_1 - \hat{V}_2 - \hat{V}_3 + \hat{V}_4$. Next,

$$\Sigma_{22} = \mathbf{E}\left(\operatorname{vec}(H)\operatorname{vec}(H)'\right) - \mathbf{E}\left(\operatorname{vec}(H)\left(\operatorname{vec}(\zeta t')\right)'\right)$$

$$- \mathbf{E}\left(\operatorname{vec}(\zeta t')\operatorname{vec}(H)'\right) + \mathbf{E}\left(\operatorname{vec}(\zeta t')\operatorname{vec}(\zeta t')'\right)$$

$$:= W_1 - W_2 - W_2' + W_3.$$

We have $\hat{W}_1 = \overline{\operatorname{vec}(H)\operatorname{vec}(H)'}$. The entries of W_2, W_3, and W_4 can be transformed to the weighted sums of the values ξ^r, and we construct the corresponding approximations by replacing ξ^r by $\overline{t_r(x)}$. Then we set $\hat{\Sigma}_{22} = \hat{W}_1 - \hat{W}_2 - \hat{W}_2' + \hat{W}_3$.

Thus, we have described the way to construct the approximation of the covariance matrix in (17), i.e.,

$$\begin{pmatrix} \hat{\Sigma}_{11} & \hat{\Sigma}_{12} \\ \hat{\Sigma}_{12}' & \hat{\Sigma}_{22} \end{pmatrix},$$

and A_n^2 in (17) is well defined.

A test of the score type is then defined as follows:

$$T_n^2 := \frac{n}{A_n^2}\left[\frac{1}{n}\sum_{i=1}^{n}\left(y_i - s(x_i)'\hat{\beta}\right)e^{\lambda x_i}\right]^2. \tag{18}$$

Denote

$$\zeta_{\tilde{o}} := (\xi, \dots, \xi^k)',$$

$$s_{\tilde{o}} := (s_1, \dots, s_k)',$$

$$H_{\tilde{o}} := \begin{bmatrix} t_1 & t_2 & \cdots & t_{k+1} \\ t_2 & t_3 & \cdots & t_{k+2} \\ \vdots & \vdots & \ddots & \vdots \\ t_k & t_{k+1} & \cdots & t_{2k} \end{bmatrix},$$

$$\beta_{\tilde{o}} := (\beta_1, \dots, \beta_k)'.$$

Then the vector η presented in (16) is equal to $\eta = [0, \eta'_{\tilde{o}}]'$, where

$$\eta_{\tilde{o}} := (\zeta_{\tilde{o}} - s_{\tilde{o}})e^{\lambda x} + (H_{\tilde{o}} - \zeta_{\tilde{o}} t')S^{-1}\mathcal{M}(\zeta e^{\lambda\xi})\mu_0,$$

and we can rewrite expression (15) in the form

$$\sqrt{n}T_{n0} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} \varepsilon_i\left(e^{\lambda x_i} - t(x_i)' f\right) + \beta'_{\tilde{o}}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_{\tilde{o},i} + o_p(1). \tag{19}$$

The covariance matrix of $\eta_{\tilde{o}}$ can be expressed as follows (cf. σ_T^2 in Lemma 3):

$$\Sigma := \operatorname{cov}(\eta_{\tilde{o}}) = \operatorname{cov}\left[I_k(\zeta_{\tilde{o}} - s_{\tilde{o}})e^{\lambda x} + (f' \otimes I_k)\operatorname{vec}(H_{\tilde{o}} - \zeta_{\tilde{o}} t')\right].$$

Hereafter, $I_k \in \mathbf{R}^{k\times k}$ is the unit matrix and

$$\Sigma = [\, I_k, f' \otimes I_k \,]\operatorname{cov}\begin{pmatrix} (\zeta_{\tilde{o}} - s_{\tilde{o}})e^{\lambda x} \\ \operatorname{vec}(H_{\tilde{o}} - \zeta_{\tilde{o}} t') \end{pmatrix}\begin{bmatrix} I_k \\ f \otimes I_k \end{bmatrix}. \tag{20}$$

Since $T_n^2 = (\sqrt{n}T_{n0})^2 / A_n^2$ and A_n^2 is a consistent estimator of σ_T^2, we obtain by Lemma 3 the following result:

Theorem 1. *Suppose that the conditions of Lemma 3 are satisfied. Also assume that one of the following two conditions is satisfied:*

(xi) $\sigma_\varepsilon^2 \mathcal{M}\left[\mathbf{E}(e^{\lambda x} - t'f)^2\right] \neq 0$;

(xii) $\beta_k \neq 0$ *and the matrix*

$$\Phi := \mathcal{M}\left(\operatorname{cov}\begin{pmatrix} (\zeta_{\tilde{o}} - s_{\tilde{o}})e^{\lambda x} \\ \operatorname{vec}(H_{\tilde{o}} - \zeta_{\tilde{o}} t') \end{pmatrix}\right)$$

is nonsingular.

Then, under the hypothesis H_0, *we have* $T_n^2 \xrightarrow{d} \chi_1^2$.

4. Power Properties of the Test

Consider the following sequence of models indexed by n:

$$H_{1\,n}:\ \ y_i = \zeta_i'\beta + \frac{1}{\sqrt{n}}g(\xi_i) + \varepsilon_i, \quad x_i = \xi_i + \delta_i, \quad i = 1, \ldots, n, \tag{21}$$

where $g: \mathbf{R} \to \mathbf{R}$ is a given function. We list the restrictions on g.

(xiii) $\mathcal{M}(ge^{\lambda\xi})$ and $\mathcal{M}(g\xi^r)$ exist, $r = 0, 1, \ldots, k$.

(xiv) $\frac{1}{n}\overline{g^2(1+\xi^{2k}+e^{2\lambda\xi})} \to 0$ as $n \to \infty$.

Under $H_{1\,n}$, for certain C we have

$$\frac{1}{\sqrt{n}A_n}\sum_{i=1}^{n}\left(y_i - s(x_i)'\hat{\beta}\right)e^{\lambda x_i} \xrightarrow{d} N(C, 1). \tag{22}$$

Relation (22) yields the following result:

Theorem 2. *Suppose that the conditions of Theorem 1 and conditions (xiii) and (xiv) are satisfied. Then, under $H_{1\,n}$, we have*

$$T_n^2 \xrightarrow{d} \chi_1^2(C), \tag{23}$$

where

$$C := \frac{\mu_0}{\sigma_0}\left(\mathcal{M}(ge^{\lambda\xi}) - \mathcal{M}(g\zeta')S^{-1}\mathcal{M}(\zeta e^{\lambda\xi})\right), \tag{24}$$

and $\chi_1^2(C)$ is a noncentral chi-squared random variable with one degree of freedom and noncentrality C.

Remark 2. To make the procedure more stable, it is better to use the test statistics T_{n0} and T_n^2 given in (14) and (18) with small sample modifications $\hat{\beta}_M$ and $\hat{\sigma}^2_{\varepsilon,M}$ instead of $\hat{\beta}$ and $\hat{\sigma}^2_{\varepsilon}$. Cheng et al. [3] showed that, for a small sample size, $\hat{\beta}_M$ provides better approximation of β than the ALS estimator, while $\sqrt{n}\left(\hat{\beta}_M - \hat{\beta}\right) \xrightarrow{P} 0$. The latter relation implies that Theorems 1 and 2 remain valid for the modified test $T^2_{n,M}$ as well. In the tests considered below (Sec. 5.2), it is also preferable for stability reasons to incorporate $\hat{\beta}_M$ and $\hat{\sigma}^2_{\varepsilon,M}$ instead of $\hat{\beta}$ and $\hat{\sigma}^2_{\varepsilon}$.

From Theorem 1, we can determine the asymptotic critical values by chi-squared distribution. By Theorem 2, the asymptotic power of T_n^2 against the local alternative (21) is $2 - \Phi\left(\lambda_{\alpha/2} - C\right) - \Phi\left(\lambda_{\alpha/2} + C\right)$, where Φ is the standard normal d.f. and $\lambda_{\alpha/2}$ is the quantile of normal law. The asymptotic power is an increasing function of $|C|$. Therefore, the larger $|C|$, the more powerful test we will have.

First, we discuss for what g the power is the largest. Till the end of this section, we assume for simplicity that ξ_i are IID random variables independent of $\{\varepsilon_i, \delta_i, \ i = 1, 2, \dots\}$. Then

$$\frac{\sigma_T C}{\mu_0} = \mathbf{E}(g e^{\lambda\xi}) - \mathbf{E}(g\zeta')(\mathbf{E}\zeta\zeta')^{-1}\mathbf{E}(\zeta e^{\lambda\xi}) = \mathbf{E}\big(g h_\lambda(\xi)\big), \tag{25}$$

where h_λ comes from the orthogonal expansion $e^{\lambda\xi} = p(\xi) + h_\lambda(\xi)$ with a polynomial $p(\xi)$, $\deg p(\xi) \le k$, and $\mathbf{E}\big(\xi^r h_\lambda(\xi)\big) = 0$, $r = 0, \dots, k$. The ratio $C^2 / \| g(\xi) \|_{L_2}^2$ is maximal if $g(\xi)$ is proportional to $h_\lambda(\xi)$, say, $g(\xi) = h_\lambda(\xi)$. Since the moments of ε are unknown, we give a consistent estimator for h_λ. We have

$$h_\lambda(\xi) = e^{\lambda\xi} - (\mathbf{E}\zeta\zeta')^{-1/2}\mathbf{E}(e^{\lambda\xi}\zeta')\zeta(\xi)(\mathbf{E}\zeta\zeta')^{-1/2},$$

and the desired approximation is given by

$$\hat{h}_\lambda(\xi) = e^{\lambda\xi} - \big(\overline{H}\big)^{-1/2}\frac{1}{\mu_0}\overline{e^{\lambda x} s(x)'}\zeta(\xi)\big(\overline{H}\big)^{-1/2}.$$

This is, up to a constant factor, an asymptotically optimal choice of $g(\xi)$ in a local alternative (21), when the weight function $w(x) = e^{\lambda x}$ is fixed.

Now consider the opposite problem. Assuming that g is fixed, we want to optimally choose the exponent λ. The function $C = C(\lambda)$ is given in (24), and we have to maximize $C^2(\lambda)$ in the domain $\lambda \in (-\infty, 0) \cup (0, \infty)$ (provided that all exponential moments of ξ and δ exist). This is a nonlinear optimization problem, and it can be solved numerically. Of course, one has to incorporate the approximations for $C^2(\lambda)$ constructed by the given data.

Consider the border case $\lambda \to 0$. Then, under the regularity conditions, the right-hand side of (25) is equal to

$$\Psi = \sum_{j=0}^{k+1} \frac{\lambda^j}{j!}\Big[\mathbf{E}(g\xi^j) - \mathbf{E}(g\zeta')(\mathbf{E}\zeta\zeta')^{-1}\mathbf{E}(\zeta\xi^j)\Big] + o(\lambda^{k+1}).$$

However, for $j = 0, 1, \dots, k$, the expression in brackets is equal to $\mathbf{E}(g\xi^j) - \mathbf{E}(g\zeta')e_j = 0$, where $e_j \in \mathbf{R}^{(k+1)\times 1}$, the jth component of e_j is equal to 1, and the other components (from the zero component to the kth component) are equal to 0. Assume that $\mathbf{E}\big(g\xi_\perp^{k+1}\big) \ne 0$, where $\xi_\perp^{k+1}$ is the orthogonal component from the expansion of ξ^{k+1} with respect to $L_{k+1} := \operatorname{span}(1, \xi, \dots, \xi^k)$ in the space of random variables L_2. Then

$$\kappa := \mathbf{E}(g\xi^{k+1}) - \mathbf{E}(g\zeta')^{-1}(\mathbf{E}\zeta\zeta')^{-1}\mathbf{E}(\zeta\xi^{k+1}) \ne 0,$$

and $\Psi = \kappa\lambda^{k+1}/(k+1)! + o(\lambda^{k+1})$ as $\lambda \to 0$. We now investigate the behavior of σ_T^2 as $\lambda \to 0$. For the value (16), we have the following expansion for small λ:

$$\eta_i \approx (\zeta_i - s(x_i))\lambda x_i + (H_i - \zeta_i t(x_i)')S^{-1}\lambda \mathbf{E}(\zeta x);$$

it has the order λ. Then $\sigma_T^2 = \operatorname{var}(\beta'\xi_i) + \sigma_\varepsilon^2 \mathbf{E}(e^{\lambda x} - t'f)^2$ has the order λ^p, $p = 0$ or 2, and relation (25) implies that $C^2(\lambda)$ has the order $\lambda^{2k+2}/\lambda^p = \lambda^{2k+2-p}$. Therefore, $\lim_{\lambda\to 0} C^2(\lambda) = 0$, and, for small λ, the test has the trivial power.

It is also possible to show that, for a polynomial $g(\xi)$ with $\deg g \geq k+1$ and for Gaussian ξ, we have $\lim_{\lambda\to 0} C^2(\lambda) = 0$. In this case, there exists an optimal $\lambda \in (-\infty, 0)\times(0, +\infty)$.

Finally, assume that ξ is normal and $g(\xi) = e^{\lambda_0 \xi}$. Then we observe a kind of a resonance effect, and, for large λ_0, the optimal exponent $\lambda_{\text{opt}} \approx \lambda_0$.

In some cases, the conditions of Theorem 2 can be satisfied only for $|\lambda| \leq \text{const}$. Then λ_{opt} should be sought on this finite interval.

5. Particular Case

We specify results in two important cases.

5.1. Linear Functional Model. We set $k = 1$ in model (1), (2). Thus, we consider the linear model

$$y_i = \beta_0 + \beta_1 \xi_i + \varepsilon_i, \quad x_i = \xi_i + \delta_i, \quad i = 1, \ldots, n,$$

where ξ_i are nonrandom. In this case, $s(x) = (1, x - \mu_1/\mu_0)'$ and $T_{n0} = \overline{\left(y - \hat\beta_0 - (x - \mu_1/\mu_0)\hat\beta_1\right)e^{\lambda x}}$, where $\hat\beta = \left(\hat\beta_0, \hat\beta_1\right)'$ is the ALS estimator. In (15), we now have

$$\sqrt{n}\,T_{n0} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} \varepsilon_i\left(e^{\lambda x_i} - t(x_i)' f\right) + \beta_1 \frac{1}{\sqrt{n}}\sum_{i=1}^{n} \eta_{\delta,i} + o_p(1)$$

and

$$\eta_\delta = \left(\frac{\mu_1}{\mu_0} - \delta\right)e^{\lambda x} + \left(\delta, \xi\delta + \delta^2 - \sigma_\delta^2\right) f,$$

$$f = \begin{pmatrix} 1 & \mathcal{M}(\xi) \\ \mathcal{M}(\xi) & \mathcal{M}(\xi^2) \end{pmatrix}^{-1} \mathcal{M}(\zeta e^{\lambda\xi})\mu_0, \quad \zeta = (1, \xi)'.$$

Then [see (20)]

$$\operatorname{var}(\eta_\delta) = [1, f']\operatorname{cov}\left(\left(\frac{\mu_1}{\mu_0} - \delta\right)e^{\lambda x}, \delta, \xi\delta + \delta^2 - \sigma_\delta^2\right)[1, f']'. \tag{26}$$

A consistent estimator of f is given by

$$\hat{f} = \begin{pmatrix} 1 & \overline{x} \\ \overline{x} & \overline{x^2} - \sigma_\delta^2 \end{pmatrix}^{-1} \left(\overline{e^{\lambda x}}, \overline{\left(x - \frac{\mu_1}{\mu_0} \right) e^{\lambda x}} \right)'.$$

A procedure for constructing the approximation $\widehat{\mathrm{cov}}$ of the covariance matrix in (26) is described in Sec. 3. For this purpose, one has to approximate $\mathcal{M}(\xi)$, $\mathcal{M}(\xi^2)$, $\mathcal{M}(e^{\lambda\xi})$, $\mathcal{M}(\xi e^{\lambda\xi})$, and $\mathcal{M}(\xi e^{2\lambda\xi})$. The corresponding approximations are

$$\overline{x}, \quad \overline{x^2} - \sigma_\delta^2, \quad \frac{1}{\mu_0}\overline{e^{\lambda x}}, \quad \frac{1}{\mu_0}\overline{\left(x - \frac{\mu_1}{\mu_0} \right) e^{\lambda x}}, \quad \text{and} \quad \frac{\overline{e^{\lambda x}}}{\mathbf{E} e^{2\lambda\delta}}.$$

Then it is easy to define A_n^2 that approximates σ_T^2:

$$A_n^2 = \left(\hat{\beta}_1\right)^2 \left[1, \hat{f}'\right] \widehat{\mathrm{cov}} \left[1, \hat{f}'\right] + \hat{\sigma}_\varepsilon^2 \overline{\left(e^{\lambda x} - t(x)' \hat{f} \right)^2}.$$

The proposed test statistic is given by

$$T_n^2 := \frac{n}{A_n^2} \overline{\left[\left(y - \hat{\beta}_0 - (x - \mu_1 / \mu_0) \hat{\beta}_1 \right) e^{\lambda x} \right]^2},$$

and Theorems 1 and 2 hold with $k = 1$.

For a linear model, we can compare the proposed bias correction procedure with the one from [5]. In that paper, in a structural model under the normality assumptions they had (in our notation)

$$\mathbf{E}\left[y - \beta_0 - (B\beta_1) x \,\middle|\, x \right] = 0 \quad \text{a.s.},$$

where B is a certain correcting coefficient, whence

$$\mathbf{E}\left[\left(y - \beta_0 - (B\beta_1) x \right) w(x) \right] = 0$$

for any weight function $w(x)$. Instead, in the present paper, for $w(x) = e^{\lambda x}$ and $s_1(x) = x - \mu_1 / \mu_0$, we have

$$\mathbf{E}\left[\left(y - \beta_0 - \beta_1 s_1(x) \right) w(x) \,|\, \xi \right] = 0 \quad \text{a.s.},$$

which also implies the unbiased relation

$$\mathbf{E}\left[\left(y - \beta_0 - \beta_1 s_1(x) \right) w(x) \right] = 0.$$

Our approach uses less information about the distributions in the model, and our procedure of the bias correction is totally different.

5.2. Polynomial Structural Model. In this subsection, we assume the following condition:

(xv) $\{\xi_i,\ i = 1, 2, \dots\}$ is IID sequence independent of $\{\varepsilon_i, \delta_i,\ i = 1, 2, \dots\}$, and the distribution law $\mathcal{L}(\xi)$ is unknown.

Then, by the strong law of large numbers, the limit values $\mathcal{M}(f(\xi))$ in our assumptions are equal to $\mathbf{E} f(\xi)$ a.s., provided that the expectation of $f(\xi)$ is finite. All explanations and results from Secs. 2–4 remain valid if the expectation is understood there as a conditional expectation for given ξ, while the assumptions are simplified. We give the corresponding statements. For fixed $k \ge 1$, consider the hypothesis

$$H_0' : \ \mathbf{E}[y - \beta_0 - \beta_1 \xi - \dots - \beta_k \xi^k \mid \xi] = 0 \ \text{ a.s. for some } \beta_0, \dots, \beta_k$$

versus the local alternatives

$$H_{1,n}' : \ \mathbf{E}\left[y - \beta_0 - \beta_1 \xi - \dots - \beta_k \xi^k - \frac{1}{\sqrt{n}} g(\xi) \middle| \xi \right] = 0 \ \text{ a.s. for some } \beta_0, \dots, \beta_k, \tag{27}$$

where g is a fixed Borel measurable function.

Lemmas 1′, 2′, and 3′ and Theorem 1′ presented below assume that H_0' is valid, whereas Theorem 2′ is stated under $H_{1,n}'$.

Lemma 1′. *Assume the following:*

(i)′ $\mathbf{E}\delta^{2k} < \infty$ and $\mathbf{E}\xi^{2k} < \infty$;

(iii)′ the matrix $S := \mathbf{E}(\zeta(\xi)\zeta'(\xi))$ *is nonsingular.*

Then $\hat{\beta} \to \beta$ *a.s. and* $\hat{\sigma}_\varepsilon^2 \to \sigma_\varepsilon^2$ *a.s.*

The proof follows from the central limit theorem for IID random variables.

Remark 3. Condition (iii)′ is satisfied in each of the following two cases:

(a) for a certain interval (a, b), a random variable $\xi \cdot I\ (\xi \in (a, b))$ has a positive density at (a, b);

(b) $\mathcal{L}(\xi)$ has at least $k + 1$ atoms.

Now let T_{n0} be defined by (14) with the same polynomials $s(x)$.

Lemma 2′. *Assume (i)′, (iii)′, (v), and the following:*

(ii)′ $\mathbf{E}\xi^{4k} < \infty$;

(vi)′ $\mathbf{E}((1 + \xi^{2k})e^{\lambda\xi}) < \infty$.

Then

$$\sqrt{n}\,T_{n0} \;=\; \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\varepsilon_i\Big(e^{\lambda x_i} - t(x_i)'f\Big) \;+\; \beta'\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\eta_i \;+\; o_p(1), \tag{28}$$

where η_i *are independent random vectors with expectation* 0,

$$\eta_i \;=\; (\zeta_i - s(x_i))e^{\lambda x_i} \;+\; \big(H_i - \zeta_i\, t(x_i)'\big)f\,,$$

$$f \;:=\; S^{-1}\mathbf{E}(\zeta e^{\lambda\xi})\mu_0\,,$$

and S *comes from (iii)′.*

In the IID case, we do not need the Lyapunov condition for the central limit theorem, and, therefore, it is easy to take a limit in (28).

Lemma 3′. *Assume (i)′, (ii)′, (iii)′, (v), and the following conditions:*

(viii)′ $\mathbf{E}\Big[(1+\delta^{2k})e^{2\lambda\delta}\Big] \;<\; \infty;$

(ix)′ $\mathbf{E}\Big(|\xi|^{3k}\,e^{\lambda\xi}\Big) \;<\; \infty$ *and* $\mathbf{E}\Big[(1+\xi^{2k})e^{2\lambda\xi}\Big] \;<\; \infty.$

Then

$$\sqrt{n}\,T_{n0} \;\xrightarrow{\;d\;}\; N\big(0,\sigma_T^2\big),$$

where

$$\sigma_T^2 \;:=\; \sigma_\varepsilon^2\,\mathbf{E}\Big(e^{\lambda x} - t(x)'f\Big)^2 \;+\; [\,\beta', f'\otimes\beta'\,]\operatorname{cov}\begin{pmatrix}(\zeta - s)e^{\lambda x}\\ \operatorname{vec}(H) - \operatorname{vec}(\zeta t')\end{pmatrix}\begin{bmatrix}\beta\\ f\otimes\beta\end{bmatrix},$$

$f := S^{-1}\mathbf{E}(\zeta(\xi)e^{\lambda\xi})\mu_0$, *and the covariance matrix is considered for a vector that depends on random* ξ *and* $x = \xi + \delta$.

Let A_n^2 be the consistent estimator of σ_T^2 introduced in Sec. 3 and let

$$T_n^2 \;:=\; \frac{\big(\sqrt{n}\,T_{n0}\big)^2}{A_n^2}\,.$$

Theorem 1′. *Let the conditions of Lemma 3′ be satisfied. Also assume that one of the following two conditions is satisfied:*

(xi)′ $\sigma_\varepsilon^2 \mathbf{E}\left(e^{\lambda x} - t' f\right)^2 \neq 0;$

(xii)′ $\beta_k \neq 0$ *and the matrix*

$$\Phi := \operatorname{cov}\begin{pmatrix} (\zeta_{\tilde{\delta}} - s_{\tilde{\delta}})e^{\lambda x} \\ \operatorname{vec}\left(H_{\tilde{\delta}} - \zeta_{\tilde{\delta}}\, t'\right) \end{pmatrix}$$

is nonsingular; here, the covariance matrix is considered for random ξ *and* x.

Then, under H_0', *we have* $T_n^2 \xrightarrow{d} \chi_1^2$.

Remark 4. Inequality $\mathbf{E}\left(e^{\lambda x} - t' f\right)^2 \neq 0$ holds in each of the following two cases:

(a) condition (a) of Remark 3 is satisfied;

(b) $\mathcal{L}(\xi)$ has at least $k+2$ atoms.

Then, under condition (a) or (b), condition (xi)′ is satisfied, provided that $\sigma_\varepsilon^2 \neq 0$.

Remark 5. In the case $k = 1$, we have a linear structural model. Then [see (26)]

$$\Phi = \operatorname{cov}\left(\left(\frac{\mu_1}{\mu_0} - \delta\right)e^{\lambda x}, \delta, \xi\delta + \delta^2 - \sigma_\delta^2\right).$$

If $\operatorname{cov}(\delta, \delta^2, e^{\lambda\delta}, \delta e^{\lambda\delta})$ is positive definite (e.g., if δ is Gaussian), then Φ is also positive definite, and condition (xii)′ is satisfied for $\beta_k \neq 0$.

We now pass to the local alternative.

Theorem 2′. *Assume that the conditions of Theorem 1′ and the following condition are satisfied:*

(xiii)′ $\mathbf{E}\left[\left(1 + |\xi|^k + e^{\lambda\xi}\right)g(\varepsilon)\right] < \infty.$

Then, under $H_{1,n}$, *we have* $T_n^2 \xrightarrow{d} \chi_1^2(C)$, *where*

$$C := \frac{\mu_0}{\sigma_T}\left[\mathbf{E}(g e^{\lambda\xi}) - \mathbf{E}(g\zeta')S^{-1}\mathbf{E}(\zeta e^{\lambda\xi})\right],$$

and $\chi_1^2(C)$ *is a noncentral chi-squared random variable with one degree of freedom and noncentrality* C.

6. Conclusions

Using an exponential weight function, we constructed a goodness-of-fit test for a polynomial EIV model. The distributions of error can be arbitrary, but we supposed that certain moments and exponential moments of δ are finite and known. The test is based on residuals, and the bias correction is performed via the exponential moments of δ. Though the structure of the test resembles the one constructed in [5] for a structural model, our idea of the bias correction is totally different and free of the normality assumption. The test relies on the ALS estimator of the regression parameter, but, for practical use, it is better to utilize a small sample modification of the estimator, which is more stable and does not differ from the ALS estimator for large samples.

We proved that the test is asymptotically chi-squared under the null hypothesis. We introduced a local alternative by adding an additional (small) summand to the regression part and showed that, under the alternative hypothesis, the test has a noncentral chi-squared asymptotic distribution. We discussed the power of the test and two related issues: (a) the optimal choice of an alternative for a fixed weight function and (b) the optimal choice of the exponent in the weight function for a fixed local alternative.

The test is also applicable to the structural model, where the latent variable is random, but its distribution is unknown. We reformulated the results for such a model and showed that we need weaker moment conditions than in the functional case. In the structural case, the discussion of the power of the test is more transparent.

It would be interesting to develop a goodness-of-fit test in a structural EIV model, where the distribution of the latent variable is known (say, Gaussian). This additional information should improve the power of the test. It looks reasonable in this case to incorporate the cost function or the score function of the corresponding consistent estimator, e.g., the ALS estimator or the quasilikelihood estimator. If all distributions in the model are normal, then it is better to base the test on the quasilikelihood estimator because the latter is more efficient than the ALS one (see [7] and [8]).

Appendix

Proof of Lemma 2. From (3), we have $\hat{\beta} = \overline{H}^{-1}\left(\overline{t\zeta'}\beta + \overline{t\varepsilon}\right)$. Substituting this expression into (14), we get

$$T_{n0} = \overline{\varepsilon\left(e^{\lambda x} - t'\overline{H}^{-1}\overline{se^{\lambda x}}\right)} + \beta'\left(\overline{\zeta e^{\lambda x}} - \overline{\zeta t'}\,\overline{H}^{-1}\,\overline{se^{\lambda x}}\right) := F_n + \beta' G_n. \tag{29}$$

We divide the proof into several steps.

A. First, we deal with $\sqrt{n}F_n$. Hereafter, the approximate equality $\approx$ means "up to summands of order $o_p(1)$ as $n \to \infty$."

We derive an expansion for $\overline{H}^{-1}$. For $0 \le r \le 2k$, we have

$$\mathbf{E}\left(\bar{t}_r - \overline{\xi^r}\right)^2 = \frac{1}{n^2}\sum_{i=1}^{n}\mathbf{E}\left(t_r(x_i) - \xi_i^r\right)^2 \to 0 \quad \text{as} \quad n \to \infty$$

by conditions (i) and (ii). However, $\overline{\xi^r}$ converges to $\mathcal{M}(\xi^r)$ by (ii). Therefore, $\bar{t}_r \xrightarrow{P} \mathcal{M}(\xi^r)$ and $\overline{H} \xrightarrow{P} S$ given in (iii).

Denote $\overline{\Lambda} := \overline{H} - S$, $\overline{\Lambda} \approx 0$. Then

$$\overline{H}^{-1} = \left(I_{k+1} + S^{-1}\overline{\Lambda}\right)^{-1} S^{-1} = S^{-1} - S^{-1}\,\overline{\Lambda}\,S^{-1} + r_n$$

and

$$\| r_n \| = \| \overline{\Lambda} \|^2 O_p(1).$$

We show that $\sqrt{n}\,\| \overline{\Lambda} \|^2 \approx 0$. Using (vii) and (ii), for $0 \le r \le 2k$ we consider

$$\mathbf{E}\left[\left(\bar{t}_r - \mathcal{M}\left(\xi^r\right)\right)^2\right] = \frac{1}{n^2}\sum_{i=1}^{n} \mathbf{E}\left(t_r(x_i) - \xi_i^r\right)^2 + \left(\overline{\xi^r} - \mathcal{M}\left(\xi^r\right)\right)^2 = \frac{O(1)}{n} + \frac{o(1)}{\sqrt{n}}.$$

Then $\sqrt{n}\left(\overline{t_r} - \mathcal{M}\left(\xi^r\right)\right)^2 \approx 0$, which implies that $\sqrt{n}\,\| \overline{\Lambda} \|^2 \approx 0$ and $\|r_n\| = o_p(1) / \sqrt{n}$. Then

$$\sqrt{n}\,\overline{\varepsilon t'}\,\overline{H}^{-1}\,\overline{se^{\lambda x}} \approx \sqrt{n}\,\overline{\varepsilon t'}\,S^{-1}\,\overline{se^{\lambda x}} - \sqrt{n}\,\overline{\varepsilon t'}\,S^{-1}\,\overline{\Lambda}\,\overline{se^{\lambda x}}.$$

However, the last summand converges to 0 in probability because $\sqrt{n}\,\overline{\varepsilon t'} = O_p(1)$, $\overline{se^{\lambda x}} = O_p(1)$, and $\overline{\Lambda} \approx 0$. By the definition of $s(x)$, we have

$$p\lim_{n\to\infty} \overline{s(x)e^{\lambda x}} = \lim_{n\to\infty} \overline{\mathbf{E}(se^{\lambda x})} = \lim_{n\to\infty} \overline{\zeta e^{\lambda \xi}}\,\mu_0.$$

Therefore,

$$\sqrt{n}\,\overline{\varepsilon t'}\,\overline{H}^{-1}\,\overline{se^{\lambda x}} \approx \sqrt{n}\,\overline{\varepsilon t'}\,S^{-1}\mathcal{M}(\zeta e^{\lambda \xi})\mu_0.$$

Thus,

$$\sqrt{n}\,F_n \approx \sqrt{n}\,\overline{\varepsilon\left(e^{\lambda x} - t' S^{-1}\mathcal{M}(\zeta e^{\lambda \xi})\mu_0\right)} = \sqrt{n}\,\overline{\varepsilon\left(e^{\lambda x} - t' f\right)}. \tag{30}$$

B. Consider $\sqrt{n}\,G_n$. We have

$$\sqrt{n}\,G_n \approx \sqrt{n}\left(\overline{\zeta e^{\lambda x}} - \overline{\zeta t'}\,S^{-1}\,\overline{se^{\lambda x}} + \overline{\zeta t'}\,S^{-1}\,\overline{\Lambda}\,S^{-1}\,\overline{se^{\lambda x}}\right).$$

But, similarly to $n^{1/4}\| \overline{\Lambda} \| \approx 0$, it is easy to show that $\sqrt[4]{n}\left(\overline{\zeta t'} - S\right) \approx 0$; we also have $\overline{se^{\lambda x}} = O_p(1)$. Then

$$\sqrt{n}\,\overline{\zeta t'}\,S^{-1}\overline{\Lambda}\,S^{-1}\,\overline{se^{\lambda x}} \approx \sqrt{n}\,\overline{\Lambda}\,S^{-1}\,\overline{se^{\lambda x}} = \sqrt{n}\left(\overline{H}\,S^{-1}\,\overline{se^{\lambda x}} - \overline{se^{\lambda x}}\right),$$

$$\sqrt{n}\,G_n \approx \sqrt{n}\,\overline{(\zeta - s)e^{\lambda x} + (H - \zeta t')S^{-1}\overline{se^{\lambda x}}}.$$

Again, $\sqrt{n}\,\overline{H - \zeta t'} = O_p(1)$ because $\mathbf{E}(H - \zeta t') = \zeta\zeta' - \zeta'\zeta = 0$ and $\overline{se^{\lambda x}} \xrightarrow{P} \mathcal{M}(\mathbf{E}\,s\,e^{\lambda x}) = \mathcal{M}(\zeta e^{\lambda \xi})\mu_0$.

Therefore,

$$\sqrt{n}\,G_n \;\approx\; \sqrt{n}\,\overline{(\zeta - s)e^{\lambda x} + (H - \zeta t')\,f}\,. \tag{31}$$

Using (29) to (31), we obtain representation (15), (16).

Proof of Lemma 3. By Lemma 2, we have

$$\sqrt{n}\,T_{n0} \;\approx\; \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_i, \tag{32}$$

where

$$Z_i \;:=\; \varepsilon_i\left(e^{\lambda x_i} - t(x_i)'f\right) \;+\; \beta'\left[(\zeta_i - s(x_i))e^{\lambda x_i} + \left(H_i - \zeta_i\, t(x_i)'\right)f\,\right]$$

are independent random vectors with expectation 0. We apply the central limit theorem in the Lyapunov form to the right-hand side of (32). We have to check the following two conditions:

(a) $\lim\limits_{n\to\infty}\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}\mathbf{E}Z_i^2 \;=\; \sigma_T^2;$

(b) $\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}\mathbf{E}|Z_i|^{2+\alpha} \;\le\; \text{const}$ for fixed $\alpha > 0$.

For Z_i, we have the representation

$$Z_i \;=\; \varepsilon_i\left(e^{\lambda x_i} - t(x_i)'f\right) \;+\; [\,\beta', f'\otimes\beta'\,]\begin{bmatrix}(\zeta_i - s(x_i))e^{\lambda x_i}\\ \operatorname{vec}(H_i) - \operatorname{vec}\left(\zeta_i\, t(x_i)'\right)\end{bmatrix}.$$

Then

$$\mathbf{E}Z_i^2 \;=\; \sigma_\varepsilon^2\,\mathbf{E}\left(e^{\lambda x_i} - t(x_i)'f\right)^2 \;+\; [\,\beta', f'\otimes\beta'\,]\operatorname{cov}\begin{bmatrix}(\zeta_i - s(x_i))e^{\lambda x_i}\\ \operatorname{vec}(H_i) - \operatorname{vec}\left(\zeta_i\, t(x_i)'\right)\end{bmatrix}\begin{bmatrix}\beta\\ f\otimes\beta\end{bmatrix},$$

whence condition (a) follows. The limit exists due to conditions (i), (ii), (vii), and (ix).

The boundedness (b) takes place due to (viii) and (x). Then, by the central limit theorem, we get

$$n^{-1/2}\sum_{i=1}^{n} Z_i \;\xrightarrow{\;d\;}\; N\left(0, \sigma_T^2\right),$$

and Lemma 3 follows from relation (32) by virtue of the Slutsky lemma.

Proof of Lemma 1. The additional conditions of Theorem 1 guarantee that σ_T^2 given in Lemma 3 is positive. Then $\left(\sqrt{n}\,T_{n0}\right)^2/\sigma_T^2 \xrightarrow{d} \chi_1^2$. However, A_n^2 is a consistent estimator of σ_T^2, and, therefore $T_n^2 = \left(\sqrt{n}\,T_{n0}\right)^2/A_n^2 \xrightarrow{d} \chi_1^2$ under the hypothesis H_0.

Proof of Theorem 2. Now assume that $H_{1,n}$ is valid. Then

$$\hat{\beta} = \overline{H}^{-1}\left(\overline{t\zeta'}\beta + \overline{t\varepsilon}\right) + \overline{H}^{-1}\frac{1}{\sqrt{n}}\overline{tg(\xi)}. \tag{33}$$

However, $\overline{H}^{-1}\left(\overline{t\zeta'}\beta + \overline{t\varepsilon}\right) \xrightarrow{P} \beta$ as the ALE estimator of β under H_0, $\overline{H}^{-1} = O_p(1)$, and

$$\mathbf{E}\left(\frac{1}{\sqrt{n}}\overline{tg(\xi)}\right)^2 = \frac{1}{n}\left(\overline{\xi g(\xi)}\right)^2 + \frac{1}{n}\overline{g^2(\xi)\mathbf{E}(t-\xi)^2},$$

which tends to 0 by virtue of (xiii) and (xiv). Therefore, relation (33) implies that $\hat{\beta} \xrightarrow{P} \beta$ under $H_{1,n}$ as well.

Now let $y = \tilde{y} + g(x)/\sqrt{n}$, where $\tilde{y} := y\big|_{H_0}$, i.e., $\tilde{y}$ is the value of y under H_0. We substitute it and (33) into (14) and note that

$$\sqrt{n}\,T_{n0}\Big|_{H_{1,n}} = \sqrt{n}\,T_{n0}\Big|_{H_0} + \left(\overline{ge^{\lambda x}} - \overline{gt'}\,\overline{H}^{-1}\,\overline{se^{\lambda x}}\right), \tag{34}$$

where $T_{n0}\big|_{H_{1,n}}$ and $T_{n0}\big|_{H_0}$ are the values of T_{n0} under $H_{1,n}$ and H_0, respectively. However, due to assumptions (xiv) and (xv), as $n \to \infty$ we have

$$\overline{ge^{\lambda x}} \approx \overline{g(\xi)\mathbf{E}e^{\lambda x}} = \mu_0\,\overline{g\xi e^{\lambda\xi}} \to \mu_0\mathcal{M}(g\,e^{\lambda\xi})$$

and $\overline{ge^{\lambda x}} \xrightarrow{P} \mu_0\mathcal{M}(\zeta e^{\lambda\xi})$ (see the proof of Lemma 2), $\overline{H}^{-1} \xrightarrow{P} S^{-1}$, and $\overline{gt'} \approx \overline{g\mathbf{E}t'} = \overline{g\zeta'} \to \mathcal{M}(g\zeta')$. Therefore, relation (34) yields

$$\sqrt{n}\,T_{n0}\Big|_{H_{1,n}} \xrightarrow{d} N\left(C_1, \sigma_T^2\right), \tag{35}$$

where $C_1 = \mu_0\left[\mathcal{M}(ge^{\lambda\xi}) - \mathcal{M}(g\zeta')S^{-1}\mathcal{M}(\zeta e^{\lambda\xi})\right]$.

The conditions of Theorem 1 are satisfied, and, from the proof of Theorem 1, we have $\sigma_T \neq 0$. Therefore, relation (35) yields

$$\frac{\left(\sqrt{n}\,T_{n0}\Big|_{H_{1,n}}\right)^2}{\sigma_T^2} \xrightarrow{d} \chi_1^2\left(\frac{C_\infty}{\sigma_T}\right). \tag{36}$$

Further, due to (xiii) and (xiv), we have $\hat{\sigma}_\varepsilon^2 \xrightarrow{P} \sigma_\varepsilon^2$ under $H_{1,n}$ as well, and, therefore the estimator A_n^2

constructed in (17) converges in probability to σ_T^2 under $H_{1,n}$ as well. Then it follows from relation (36) that $T_n^2\big|_{H_{1,n}} \xrightarrow{d} \chi_1^2(C)$.

Proof of the Statement in Remark 3. We must show that $1, \xi, \ldots, \xi^k$ are linearly independent in L^2. Assume that, for certain constants $a_0, \ldots, a_k$, we have $a_0 + a_1\xi + \ldots + a_k\xi^k = 0$ a.s. In case (a), we have $a_0 + a_1 u + \ldots + a_k u^k = 0$ for $u \in (a, b)$ almost everywhere with respect to the Lebesgue measure, which implies that $a_0 = a_1 = \ldots = a_k = 0$, and $1, \xi, \ldots, \xi^k$ are linearly independent in L^2.

In case (b), we have $a_0 + a_1 u_j + \ldots + a_k (u_j)^k = 0$, $j = 1, 2, \ldots, k+1$, where $u_1, \ldots, u_{k+1}$ are the atoms of the distribution $\mathcal{L}(\xi)$. This again yields $a_0 = a_1 = \ldots = a_k = 0$.

Proof of Lemma 2′. It follows the line of the proof of Lemma 2. We only point out the expansion of $\overline{H}^{-1}$. In this case, for IID ξ_i, we have $\mathbf{E}(H) = S$, and, by virtue of (i) and (ii)′,

$$\overline{\Lambda} = \overline{H} - \mathbf{E}(\overline{H}) = \frac{O_p(1)}{\sqrt{n}}.$$

Then $\sqrt{n}\|\overline{\Lambda}\|^2 \approx 0$ and $\overline{H}^{-1} \approx \sqrt{n}\left(S^{-1} - S^{-1}\overline{\Lambda}S^{-1}\right)$. The other computations are performed by analogy with the proof of Lemma 2.

Proof of Lemma 3′. We apply the central limit theorem for and IID sequence to the sums on the right-hand side of (28). Conditions (viii)′ and (ix)′ guarantee the existence of the corresponding second moments.

Proof of Theorem 1′. Under condition (xi)′ or (xii)′, the asymptotic variance σ_T^2 is positive, and Theorem 1′ can be proved by analogy with Theorem 1.

Proof of the Statement in Remark 4. We prove this statement by contradiction. Let $e^{\lambda x} - t(x)'f = 0$ a.s. Then $0 = \mathbf{E}\left[e^{\lambda x} - t(x)'f \,|\, \xi\right] = \mu_0 e^{\lambda\xi} - \zeta(\xi)'f$ a.s. and

$$e^{\lambda\xi} = \sum_{i=0}^{k} \frac{f_i}{\mu_0}\xi^i \quad \text{a.s.}$$

In both case (a) and case (b), we establish that

$$e^{\lambda u} = \sum_{i=0}^{k} \frac{f_i}{\mu_0}u^i$$

for at least $k+2$ different values of u. Then there exists a point u_0 at which one has either

$$\frac{d^{k+1}e^{\lambda u}}{du^{k+1}} = \frac{d^{k+1}}{du^{k+1}}\left(\sum_{i=0}^{k}\frac{f_i}{\mu_0}u^i\right)$$

or $\lambda^{k+1}e^{\lambda u_0} = 0$. However, this is impossible for $\lambda \neq 0$.

Proof of Theorem 2′. The ALS estimators $\hat{\beta}$ and $\hat{\sigma}_\varepsilon^2$ are strongly consistent estimators of β and σ_ε^2, respectively, under $H_{1,n}$ as well, and $A_n^2 \to \sigma_T^2$ a.s. under $H_{1,n}$. Relation (34) is now true. By virtue of condition (xiii)′ and the strong law of large numbers, we conclude that, almost surely,

$$\lim_{n\to\infty}\left(\overline{ge^{\lambda x}} - \overline{gt'}\,\overline{H}^{-1}\,\overline{se^{\lambda x}}\right) = \mathbf{E}(ge^{\lambda x}) - \mathbf{E}(gt')S^{-1}\mathbf{E}se^{\lambda x} := C_1$$

and

$$T_n^2\Big|_{H_{1,n}} = \frac{\left(\sqrt{n}T_{n0}\right)^2}{A_n^2}\Big|_{H_{1,n}} \xrightarrow{d} \chi_1^2\left(\frac{C_1}{\sigma_T}\right) = \chi_1^2(C),$$

where C is defined in Theorem 2′.

The authors are grateful to Prof. H. Schneeweiss (Munich, Germany) for fruitful discussions.

The research was partially realized in the course of a visit of A. Kukush to the Institute of Statistical Science of Academia Sinica (Taipei, Republic of China). This work was partially supported by a postdoctoral fellowship of the Belgian Office for Scientific, Technical, and Cultural Affairs.

REFERENCES

1. C.-L. Cheng and H. Schneeweiss, "Polynomial regression with errors in the variables," *J. R. Statist. Soc. B,* **60**, 189–199 (1998).
2. R. J. Carroll, D. Ruppert, and L. A. Stefanski, *Measurement Error in Nonlinear Models,* Chapman and Hall, London (1995).
3. C.-L. Cheng, H. Schneeweiss, and M. Thamerus, "A small sample estimator for a polynomial regression with errors in the variables," *J. R. Statist. Soc. B,* **62**, 699–709 (2000).
4. C.-L. Cheng and H. Schneeweiss, "On the polynomial measurement error model," in: S. van Huffel and P. Lemmerling (editors), *Total Least Squares and Errors-in-Variables Modeling,* Kluwer, Dordrecht (2002), pp. 131–143.
5. L. Zhu, H. Cui, and K. W. Ng, "Testing lack-of-fit for linear errors-in-variables model," *Acta Appl. Math.* (to appear).
6. H. P. Rosenthal, "On the subspaces of L^p $(p > 2)$ spanned by sequences of independent random variables," *Isr. J. Math.,* **8**, 273–303 (1970).
7. H. Schneeweiss and T. Nittner, *Estimating a Polynomial Regression with Measurement Errors in the Structural and in the Functional Case – a Comparison,* Discuss. Paper 197. Sonderforschungbereich 386, Univ. Munich, Munich (2000).
8. A. Kukush and H. Schneeweiss, "A comparison of asymptotic covariance matrices of adjusted least squares and structural least squares in error ridden polynomial regression," *J. Statist. Plan. Inference* (to appear).