UDC 519.21

**C. -L. Cheng** (Acad. Sinica, Taipei, China),

**A. G. Kukush*** (Kyiv Nat. T. Shevchenko Univ., Ukraine)

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

# ПЕРЕВІРКА АДЕКВАТНОСТІ ПОЛІНОМІАЛЬНОЇ МОДЕЛІ З ПОХИБКАМИ У ЗМІННИХ

The polynomial regression models with errors in the variables are considered. A goodness-of-fit test is constructed, which is based on adjusted least squares estimator and modifies the test, introduced by Zhu et al. for 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 the test statistics.

Побудовано процедуру перевірки адекватності поліноміальної регресійної моделі з похибками у змінних, що ґрунтується на адаптивній оцінці найменших квадратів і є модифікацією процедури, запропонованої в роботі Жу та ін. для перевірки адекватності лінійної структурної моделі з гауссовими похибками. У даній роботі не вимагається, щоб розподіл похибок був гауссовим. Запропонована процедура базується на залишкових членах і характеризується асимптотичним хі-квадрат розподілом при нульовій гіпотезі. Вивчено потужність процедури та питання про вибір показника експоненціальної вагової функції, яка використовується в процедурі.

**1. Introduction.** Cheng and Schneeweiss [1] developed an adjusted least squares (ALS) estimator of the parameters of a polynomial functional regression model with errors in the 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., Carroll et al. [2], Chapter 6). In Cheng et al. [3] a small sample modification of the ALS estimator was constructed, which shows good results in small sample and which is asymptotically equivalent to the ALS estimator. For a further discussion of related models see Cheng and Schneeweiss [1, 4].

Errors-in-variables (EIV) models are widely used in practical applications. Therefore it is relevant to develop appropriate goodness-of-fit tests. But most of the literature in the EIV context dealt with estimation rather than testing. In Zhu et al. [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 the test of 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 performed now on the basis of the exponential moments of the errors, which are supposed to be known. The test relies on the ALS estimator and its small sample modification.

Standard notations used in the paper are: $\mathbf{E}\varepsilon$ and $\operatorname{var}(\varepsilon)$ for the expectation and the variance of the random variable $\varepsilon$, $\operatorname{cov}(\xi)$ for the variance-covariance matrix of the random vector $\xi$, and $O_p(1)$ for a sequence of stochastically bounded random variables, and $o_p(1)$ for a sequence of random variables which converges to $0$ in probability.

In Section 2 we describe the model and the ALS estimator. In Section 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

* The research was partially realized while this author was visiting Institute of Statistical Science of Academia Sinica, Taipei, Republic of China. Partially supported by a postdoctoral fellowship of the Belgian office for Scientific, Technical and Cultural Affairs.

Section 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 unknown distribution. Section 6 concludes, and the proofs of the results are postponed to the Appendix.

**2. The model and the estimator.** We discuss the polynomial functional relationship

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

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

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

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 $t_r(x)$ can be expressed via the moments $\mathbf{E}\delta^l$, $l = 1, \ldots, r$, see Cheng and Schneeweiss [1]. Denote $t = t(x) := (t_0(x), t_1(x), \ldots \ldots, t_k(x))'$, 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 $\beta$ satisfies the equation

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

Hereafter the bars denote averages, i. e.,

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

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

**Lemma 1** [1]. *Assume the following*:

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

(ii) *the* $\mathcal{M}(\xi^r)$ *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* $\sigma_\varepsilon^2$, *as* $n \to \infty$.

**3. Construction of the test and bias correction.** For the response variable $y$ and the corresponding latent variable $\xi$, we consider the following hypothesis with fixed $k \geq 1$,

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

versus

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

If we want to use the residuals in a test statistics for the hypothesis $H_0$ based on observed $y$'s and $x$'s, we have to consider the expectation of the residual with $x$ in the place of $\xi$ in (5). However as discussed in Zhu et al. [5],

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

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

We perform the correction as follows. Let $w(\cdot)$ be a weight function, then $H_0$ implies an equality

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

We want to construct polynomials $s_0(x), s_1(x), \ldots, s_k(x)$ such that under $H_0$

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

where $s(x) = s = (s_0(x), \ldots, s_k(x))'$. It is possible to satisfy (8) if one choses $w(x) = e^{\lambda x}$, with fixed $\lambda \neq 0$, provided the corresponding exponential moments of $\delta$ exist. We assume the following:

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

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

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

We have $s_0(x) = 1$, $s_1(x) = x - \mu_1/\mu_0$. We look for $s_r(x)$ in the form $s_r(x) = \sum_{j=0}^{r} b_{rj} x^j$. To fulfill (9), we have the relations

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

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

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

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

$$a_{pj} = \binom{j}{p} \mu_{j-p}, \quad 0 \leq p \leq j \leq r. \tag{12}$$

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

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

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

Now, with the polynomials (13) we consider a statistic of the score type

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

Remember that the bar denotes 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 that (iv).

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

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

***Remark* 1.** If $\xi_i$ are IID random variables with $\mathbf{E}|\xi|^{8k+\alpha} < \infty$, where $\alpha > 0$ fixed, then (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 the Rosenthal moment inequality, [6],

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

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 (vii) holds a. s. by Chebyshev inequality and Borel Cantelli lemma.

Remark 1 shows that the condition (vii) is realistic, it holds a. s. for a structural polynomial EIVM if $\xi$ 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* $\eta_i$ *are independent random vectors with expectation* 0, *and*

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

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

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

Now, we introduce some more assumptions in order to apply the central limit theorem in Lyapunov form to the sum of $\eta_i$.

(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) The $\mathcal{M}(\xi^r e^{\lambda\xi})$ exists for $r = 0, 1, \ldots, 3k$, and the $\mathcal{M}(\xi^r e^{2\lambda\xi})$ exists for $r = 0, 1, \ldots, 2k$.

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

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

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

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

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

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

$Z(\xi,\delta)$ *is a vector function on* $\xi$ *and* $\delta$.

Under the conditions of Lemma 3, the approximation of $\sigma_T^2$ 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) Now,

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

and we describe the approximations $\hat\Sigma_{ij}$ to $\mathcal{M}(\Sigma_{ij}(\xi))$. We have

$$\Sigma_{11} = \mathbf{E}(s(x)s(x)'e^{2\lambda x}) + \zeta\zeta' e^{2\lambda\xi}\mathbf{E}e^{2\lambda\delta} - \zeta e^{2\lambda\xi} - \mathbf{E}(s(x)'e^{2\lambda\delta}) -$$
$$- \mathbf{E}(s(x)e^{2\lambda\delta})e^{2\lambda\xi}\zeta' := U_1 + U_2 - U_3 - U_3'.$$

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

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

$$+ \mathbf{E}\left(\operatorname{vec}(\zeta t(x)')e^{\lambda x}s(x)'\right) := V_1 - V_2 - V_3 + V_4.$$

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

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

$$- \mathbf{E}\left(\operatorname{vec}(\zeta t')\operatorname{vec}(H)'\right) + \mathbf{E}\left(\operatorname{vec}(\zeta t')\operatorname{vec}(\zeta t')'\right) := W_1 - W_2 - W_2' + W_3.$$

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 $\xi^r$, and we construct the corresponding approximations by replacing $\xi^r$ for $\overline{t_r(x)}$. Then we set $\hat{\Sigma}_{22} = \hat{W}_1 - \hat{W}_2 - \hat{W}_2' + \hat{W}_3$.

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

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

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

A test of score type is then defined by

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

Denote

$$\zeta_{\bar{o}} := (\xi, \ldots, \xi^k)',$$

$$s_{\bar{o}} := (s_1, \ldots, s_k)',$$

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

$$\beta_{\bar{o}} := (\beta_1, \ldots, \beta_k)'.$$

Then the vector $\eta$ presented in (16) equals $\eta = \left[0, \eta_{\bar{o}}'\right]'$, with

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

and we can rewrite the expression (15),

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

A covariance matrix of $\eta_{\bar{\sigma}}$ can be expressed as follows, compare with $\sigma_T^2$ in Lemma 3,

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

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

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

Since $T_n^2 = \left(\sqrt{n}T_{n0}\right)^2/A_n^2$ and $A_n^2$ is a consistent estimator of $\sigma_T^2$, we obtain by Lemma 3 the next result.

**Theorem 1.** *Let the conditions of Lemma* 3 *hold. Assume additionally one of the following two conditions*:

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

(xii) $\beta_k \neq 0$, *and a matrix*

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

*is nonsingular.*

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

**4. The power properties of the test.** Consider a sequence of models indexed by $n$ of the following form, with a given function $g: \mathbf{R} \to \mathbf{R}$:

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

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

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

From (22), the next results follows.

**Theorem 2.** *Assume the conditions of Theorem* 1 *and conditions* (xiii), (xiv). *Then under* $H_{1n}$, *we have*

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

*where*

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

*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 in the test statistics $T_{n0}$ and $T_n^2$ given in (14) and (18), with the 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 small sample size $\hat\beta_M$ provides better approximation of $\beta$ than the ALS estimator, while $\sqrt{n}\big(\hat\beta_M - \hat\beta\big) \xrightarrow{P} 0$. 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 in subsection 5.2, it is also preferable for stability reasons to incorporate $\hat\beta_M$ and $\hat\sigma^2_{\varepsilon,M}$ rather than $\hat\beta$ 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 $2 - \Phi\big(\lambda_{\alpha/2} - C\big) - \Phi\big(\lambda_{\alpha/2} + C\big)$, where $\Phi$ 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 which $g$ the power is the largest. Till the end of this section, suppose for simplicity that $\xi_i$ are IID random variables, independent from $\{\varepsilon_i, \delta_i, i = 1, 2, \ldots\}$. Then

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

where $h_\lambda$ comes from the orthogonal expansion $e^{\lambda\xi} = p(\xi) + h_\lambda(\xi)$, with a polynomial $p(\xi)$, $\deg p(\xi) \le k$, and $\mathbf{E}(\xi^r h_\lambda(\xi)) = 0$, $r = 0, \ldots, k$. The ratio $C^2/\|g(\xi)\|^2_{L_2}$ is maximal if $g(\xi)$ is proportional to $h_\lambda(\xi)$, $g(\xi) = h_\lambda(\xi)$, say. As the moments of $\varepsilon$ are unknown, we give a consistent estimator for $h_\lambda$. 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 an opposite problem. Let $g$ be fixed, and we want to choose optimally an exponent $\lambda$. 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 all the exponential moments of $\xi$ and $\delta$ 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 a border case $\lambda \to 0$. Then under regularity conditions the right-hand side of (25) equals

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

But for $j = 0, 1, \ldots, k$, the expression in brackets equals $\mathbf{E}(g\xi^j) - \mathbf{E}(g\zeta')e_j = 0$, where $e_j \in \mathbf{R}^{(k+1)\times 1}$, $j$-th component of $e_j$ equals 1, and all the rest from 0 component to $k$-th component equal 0. Assume that $\mathbf{E}\left(g\xi_{\perp}^{k+1}\right) \neq 0$, where $\xi_{\perp}^{k+1}$ is orthogonal component from the expansion of $\xi^{k+1}$ w. r. t. $L_{k+1} := \operatorname{span}(1, \xi, \ldots, \xi^k)$ in the $L_2$ space of random variables. Then

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

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

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

which has the order $\lambda$. Then $\sigma_T^2 = \operatorname{var}(\beta'\xi_i) + \sigma_\varepsilon^2\mathbf{E}(e^{\lambda x} - t'f)^2$ has the order $\lambda^p$, $p = 0$ or 2, and from (25) we obtain 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 $\lambda$ the test has trivial power.

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

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

Sometimes the conditions of Theorem 2 can be valid only for $|\lambda| \leq \text{const}$. Then $\lambda_{\text{opt}}$ has to be searched at this finite interval.

**5. Particular case.** We specify the results in two important cases.

*5.1. Linear functional model.* We set $k = 1$ in the model (1), (2). Thus we consider a 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 $\xi_i$ are nonrandom. Now, $s(x) = (1, x - \mu_1/\mu_0)'$, 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. Now, in (15) we have

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

and

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

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

Now, see (20),

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

A consistent estimator of $f$ is given by

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

The procedure for constructing the approximation $\widehat{\operatorname{cov}}$ of the covariance matrix in (26) is described in Section 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:

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

Then it is easy to define $A_n^2$, which approximates $\sigma_T^2$,

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

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 true with $k = 1$.

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

$$\mathbf{E}[y - \beta_0 - (B\beta_1)x \mid x] = 0 \text{ a. s.},$$

for certain correcting coefficient $B$, and this implied

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

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

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

which also implies the unbiased relation

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

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) $\{\xi_i,\ i = 1, 2, \ldots\}$ is IID sequence, independent from $\{\varepsilon_i, \delta_i,\ i = 1, 2, \ldots\}$, 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 will be equal to $\mathbf{E}f(\xi)$ a. s., provided the expectation of $f(\xi)$ is finite. All the explanations and results from Sections 2 – 4 are still valid if to understand there the expectation as a conditional expectation given $\xi$, while the assumptions are simplified. We give the corresponding statements. We consider the following hypothesis with fixed $k \geq 1$,

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

versus local alternatives

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

where $g$ is a fixed Borel measurable function.

The following Lemmas $1'$, $2'$, $3'$ and Theorem $1'$ assume that $H_0'$ holds true, while 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)′ holds in each of the following two cases:

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

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

Now, $T_{n0}$ is 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*

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

*where* $\eta_i$ *are independent random vectors with expectation* 0, *and*

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

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

*where* $S$ *comes from* (iii)′.

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

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

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

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

*Then* $\sqrt{n}\, T_{n0} \xrightarrow{d} N(0, \sigma_T^2)$, *where*

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

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

Let $A_n^2$ be a consistent estimator of $\sigma_T^2$ introduced in Section 3, and $T_n^2 := := (\sqrt{n}T_{n0})^2 / A_n^2$.

**Theorem 1′.** *Let conditions of Lemma 3′ hold. Suppose also that one of the following two conditions holds:*

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

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

$$\Phi := \operatorname{cov}\begin{pmatrix} (\zeta_{\bar{\alpha}} - s_{\bar{\alpha}})e^{\lambda x} \\ \operatorname{vec}(H_{\bar{\alpha}} - \zeta_{\bar{\alpha}} t') \end{pmatrix}$$

*is nonsingular, where the covariance matrix is considered for random* $\xi$ *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) when conditions a) of Remark 3 hold, or b) $\mathcal{L}(\xi)$ has at least $k+2$ atoms. And then, under a) or b), condition (xi)′ holds provided $\sigma_\varepsilon^2 \neq 0$.

***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 $\delta$ is Gaussian) then $\Phi$ is also positive definite, and condition (xii)′ holds for $\beta_k \neq 0$.

Now, we pass to the local alternative.

**Theorem 2′.** *Assume the conditions of Theorem* 1′ *and next condition*

(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*

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

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

**6. Conclusion.** 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 $\delta$ are finite and known. The test is based on residuals, and the bias correction is performed via the exponential moments of $\delta$. Though the structure of the test resembles the one constructed by Zhu et al. [5] for a structural model, our idea of the bias correction is totally different and free from normality assumption. The test relies on the ALS estimator of the regression parameter, but for practical use it is better to utilize the small sample modification of the estimator, which is more stable and does not differ from the ALS estimator for large sample.

We proved that the test is asymptotically chi-squared under 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 noncentral chi-squared asymptotic distribution. We discussed the power of the test and also two related issues: a) the optimal choice of the alternative for fixed weight function, and b) the optimal choice of the exponent in the weight function for fixed local alternative.

The test is applicable to the structural model as well, 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 has to improve the power of the test. It looks plausible 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 the distributions in the model are normal, it is better to base the test on the quasilikelihood estimator, because the latter estimator is more efficient than the ALS one, see Schneeweiss and Nittner [7] and Kukush et al. [8].

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

**Appendix.** ***Proof of Lemma 2.*** From (3) we have $\hat\beta = \overline{H}^{-1}\left(\overline{t\zeta'}\beta + \overline{t\varepsilon}\right)$. We substitute it into (14).

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

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

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

by conditions (i) and (ii). But $\overline{\xi^r}$ 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), (ii), consider for $0 \le r \le 2k$,

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

$$= \frac{O(1)}{n} + \frac{o(1)}{\sqrt{n}}.$$

Then $\sqrt{n}\left(\overline{t_r} - \mathcal{M}\left(\xi^r\right)\right)^2 \approx 0$, which implies $\sqrt{n}\|\overline{\Lambda}\|^2 \approx 0$, and $\|r_n\| = o_p(1)/\sqrt{n}$. Now, $\sqrt{n}\,\overline{\varepsilon t'}\,\overline{H}^{-1}\overline{se^{\lambda x}} \approx \sqrt{n}\,\overline{\varepsilon t'}S^{-1}\overline{se^{\lambda x}} - \sqrt{n}\,\overline{\varepsilon t'}S^{-1}\overline{\Lambda}\,\overline{se^{\lambda x}}$. But the last summands 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 have also $\overline{se^{\lambda x}} = O_p(1)$. 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 $\overline{se^{\lambda x}} \xrightarrow{P}$ $\xrightarrow{P} \mathcal{M}(\mathbf{E}se^{\lambda x}) = \mathcal{M}(\zeta e^{\lambda\xi})\mu_0$. Therefore

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

From (29) to (31) we obtain the 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

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

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) $\lim\limits_{n\to\infty}\dfrac{1}{n}\sum\limits_{i=1}^{n}\mathbf{E}Z_i^2 = \sigma_T^2.$

b) For fixed $\alpha > 0$, $\dfrac{1}{n}\sum\limits_{i=1}^{n}\mathbf{E}|Z_i|^{2+\alpha} \le \text{const}.$

For $Z_i$ we have the representation

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

Then

$$\mathbf{E} Z_i^2 = \sigma_\varepsilon^2 \mathbf{E}\left(e^{\lambda x_i} - t(x_i)' f\right)^2 +$$

$$+ [\beta', f' \otimes \beta'] \operatorname{cov}\begin{bmatrix} (\zeta_i - s(x_i)) e^{\lambda x_i} \\ \operatorname{vec}(H_i) - \operatorname{vec}(\zeta_i t(x_i)') \end{bmatrix} \begin{bmatrix} \beta \\ f \otimes \beta \end{bmatrix},$$

and from this relation the 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

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

and Lemma 3 follows from (32) by Slutsky lemma.

***Proof of Lemma* 1.** Additional conditions of Theorem 1 provide that $\sigma_T^2$ given in Lemma 3 is positive. Then $\left(\sqrt{n}\, T_{n0}\right)^2 / \sigma_T^2 \xrightarrow{d} \chi_1^2$. But $A_n^2$ is a consistent estimator of $\sigma_T^2$, therefore $T_n^2 = \left(\sqrt{n}\, T_{n0}\right)^2 / A_n^2 \xrightarrow{d} \chi_1^2$, under validity of $H_0$.

***Proof of Theorem* 2.** We assume now that $H_{1,n}$ holds. Then

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

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

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

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

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

$$\sqrt{n}\, T_{n0}\big|_{H_{1,n}} = \sqrt{n}\, T_{n0}\big|_{H_0} + \left(\overline{g e^{\lambda x}} - \overline{g t'}\, \overline{H}^{-1} \overline{s e^{\lambda x}}\right), \tag{34}$$

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. But due to assumptions (xiv) and (xv) we have, as $n \to \infty$,

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

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

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

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

Conditions of Theorem 1 hold, and from the proof of Theorem 1 we have $\sigma_T \neq 0$. Therefore (35) implies

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

Next, due to (xiii), (xiv), $\hat\sigma_\varepsilon^2 \xrightarrow{P} \sigma_\varepsilon^2$ under $H_{1,n}$ as well, therefore the estimator $A_n^2$ constructed in (17) converges in probability to $\sigma_T^2$ under $H_{1,n}$ as well. Then (36) implies $T_n^2\Big|_{H_{1,n}} \xrightarrow{d} \chi_1^2(C)$.

***Proof of statement in Remark* 3.** We have to show that $1, \xi, \ldots, \xi^k$ are linearly independent in $L^2$. Suppose 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 w. r. t. Lebesgue measure, which implies $a_0 = a_1 = \ldots = a_k = 0$, 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 $\mathcal{L}(\xi)$. And again this implies $a_0 = a_1 = \ldots = a_k = 0$.

***Proof of Lemma* 2′.** It follows the line of the proof of Lemma 2. We only point out the expansion of $\overline{H}^{-1}$. Now, for IID $\xi_i$, $\mathbf{E}(H) = S$, and by (i), (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 similarly to the proof of Lemma 2.

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

***Proof of Theorem* 1′.** Under (xi)′ or (xii)′ the asymptotic variance $\sigma_T^2$ is positive, and Theorem 1′ can be proved similarly to Theorem 1.

***Proof of statement in Remark* 4.** We prove by the contrary. Suppose that $e^{\lambda x} - t(x)'f = 0$ a. s. Then a. s. we have $0 = \mathbf{E}\left[e^{\lambda x} - t(x)'f \mid \xi\right] = \mu_0 e^{\lambda\xi} - \zeta(\xi)'f$, and

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

In both case a) or b) we obtain 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 $u_0$, such that

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

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

***Proof of Theorem* 2′.** The ALS estimators $\hat\beta$ and $\hat\sigma_\varepsilon^2$ are strongly consistent estimators of $\beta$ and $\sigma_\varepsilon^2$, respectively, under $H_{1,n}$ as well, and $A_n^2 \to \sigma_T^2$ a. s. under

$H_{1,n}$, also. Now, (34) holds true. And due to (xiii)′ by the strong law of large numbers we have a. s. that

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

and

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

with $C$ defined in Theorem 2′.


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


Received 03.10.2002