Cheng, Schneeweiss:

Note on Two Estimators for the Polynomial Regression with Errors in the Variables

Sonderforschungsbereich 386, Paper 96 (1997)

Note on Two Estimators for the Polynomial Regression with Errors in the Variables

By CHI-LUN CHENG and HANS SCHNEEWEISS

Institute of Statistical Science, Academia Sinica, Taipei, Republic of China

Institute for Statistics, University of Munich, Munich, Germany

SUMMARY: This Note generalizes two estimators of the quadratic regression with measurement errors by Fuller and Wolter and Fuller to the polynomial case.

1 Introduction

Fuller (1987) and Wolter and Fuller (1982) consider quadratic functional relationships with errors in the variables for two different cases regarding the presence (case 1) or absence (case 2) of errors in the equation. They develop estimators for the parameters of the quadratic relationship in both cases, assuming that, in case 1, the error variance of the regressor variable or, in case 2, the error variances of dependent and regressor variables be known. In case 1 the errors of dependent and regressor variables are assumed to be uncorrelated. In both cases the errors are taken to be normally distributed.

Both these estimators can be generalized to the model of a polynomial functional relationship of any degree and with correlated errors of dependent and regressor variables and with not necessarily normal errors. In case 1 the resulting estimator is seen to be the same as the one developed by Cheng and Schneeweiss (1996). These authors derived in their paper the asymptotic covariance matrix of the estimator of case 1. The same will be done in this note for the estimator of case 2.

2 Case 1: Errors in the equation

Consider a polynomial functional relationship with errors in the equation:

$$\begin{aligned} y_i &= \eta_i + \varepsilon_i = \beta_0 + \beta_1\xi_i + \beta_2\xi_i^2 + \ldots + \beta_k\xi_i^k + \varepsilon_i \\ x_i &= \xi_i + \delta_i \end{aligned} \tag{1}$$

$i = 1, \ldots, n$, where $(\delta_i, \varepsilon_i)$ are i.i.d. random errors with expectation 0 and covariance matrix

$$\Omega = \begin{pmatrix} \sigma_\delta^2 & \sigma_{\delta\varepsilon} \\ \sigma_{\delta\varepsilon} & \sigma_\varepsilon^2 \end{pmatrix}.$$

The $\xi_i, i = 1, \ldots, n$, are unobservable (latent) nonstochastic variables. The regressor error variance σ_δ^2 and the covariance $\sigma_{\delta\varepsilon}$ of regressor error δ and dependent variable error ε are assumed to be known. The variance σ_ε^2, which contains the error-in-the-equation variance component, is unknown. If the error variables are jointly normally distributed we have

$$(N) \qquad (\delta_i, \varepsilon_i) \sim N(0, \Omega).$$

It is well-known that replacing the latent variable ξ by its observable counterpart x in the polynomial relationship and estimating the parameters β_j in the resulting polynomial regression by OLS yields inconsistent estimates, Grilliches and Ringstad (1970). As a first step to remove this inconsistency, Fuller (1987) suggests to view the powers of ξ as $k+1$ different latent regressor variables, for which as their observable counterparts unbiased estimates t_r computable from the data are available, so that a linear functional relationship results:

$$\begin{aligned} y_i &= \beta_0 \xi_i^0 + \beta_1 \xi_i^1 + \ldots + \beta_k \xi_i^k + \varepsilon_i \\ t_{ri} &= \xi_i^r + e_{ri}, \quad r = 0, \ldots, k, i = 1, \ldots, n, \end{aligned}$$

where the e_{ri} are the new measurement errors, with $Ee_{ri} = 0$. Let $\zeta_i = (\xi_i^0, \ldots, \xi_i^k)'$, $\beta = (\beta_0, \ldots, \beta_k)'$, $t_i = (t_{0i}, \ldots, t_{ki})'$, $e_i = (e_{0i}, \ldots, e_{ki})'$, then the model can be written as

$$\begin{aligned} y_i &= \zeta_i' \beta + \varepsilon_i \\ t_i &= \zeta_i + e_i \end{aligned} \qquad i = 1, \ldots, n \tag{2}$$

For this linear functional relationship a consistent estimator of β can now be constructed if unbiased estimates $\hat{V}_i$ and $\hat{v}_i$ of, respectively, the covariance matrix $V_i = E(e_i e_i')$ and the covariance vector $v_i = E(e_i \varepsilon_i)$ are available. The error adjusted least squares normal equations are given by (all summations are for $i = 1, \ldots, n$)

$$\Sigma(t_i t_i' - \hat{V}_i)\hat{\beta} = \Sigma(t_i y_i - \hat{v}_i),$$

or

$$(\overline{tt'} - \bar{\hat{V}})\hat{\beta} = \overline{ty} - \bar{\hat{v}}, \tag{3}$$

where the $^-$ denotes averages over $i = 1, \ldots, n$; cf. Fuller (1987, p.212) for the quadratic relationship.

Following an idea of Chan and Mak (1985) the estimates t_{ri} of ξ_i^r are easily constructed as certain polynomials in x_i of degree r. Their coefficients depend on higher moments of δ_i up to the order of $2r$, which are assumed to be known. In case of (N) only σ_δ^2 needs to be known, and the t_{ri} can be computed by the recursive relation $t_{r+1,i} = x_i t_{ri} - \sigma_\delta^2 r t_{r-1,i}$ with $t_0 = t_{-1} = 1$, cf. Cheng and Schneeweiss (1996).

The covariance matrix V_i is given by

$$\begin{aligned} V_i &= E(t_i - \zeta_i)(t_i - \zeta_i)' \\ &= E(t_i t_i') - \zeta_i \zeta_i', \end{aligned} \tag{4}$$

The elements of $\zeta_i\zeta_i'$ are powers of ξ_i and can therefore be estimated by the variables t_{ri}. Let

$$H_i = \begin{pmatrix} t_{0i} & t_{1i} & \dots & t_{ki} \\ t_{1i} & t_{2i} & \dots & t_{k+1,i} \\ \vdots & \vdots & & \vdots \\ t_{ki} & t_{k+1,i} & \dots & t_{2k,i} \end{pmatrix},$$

then $EH_i = \zeta_i\zeta_i'$. Thus, an unbiased estimate of V_i is given by

$$\hat{V}_i = t_i t_i' - H_i.$$

Similarly,

$$v_i = E\{(t_i - \zeta_i)\varepsilon_i\} = E(t_i\varepsilon_i).$$

Cheng and Schneeweiss (1996) derive an unbiased estimate of $E(t_{ri}\varepsilon_i)$ in terms of a linear combination of the t_{ri}, the coefficients of which depend on $E\delta_i^l$ and $E(\delta_i^l\varepsilon_i)$ only, $l = 1, \dots, r$, and which they denote by $\hat{E}(t_{ri}\varepsilon_i)$. Thus $\hat{v}_i = \hat{E}(t_i\varepsilon_i)$. In case of (N), $\hat{E}(t_{ri}\varepsilon_i) = \sigma_{\delta\varepsilon}\ r t_{r-1,i}$.

The normal equations (1) can now be written as

$$\bar{H}\hat{\beta} = \bar{h}$$

with $\bar{H} = \frac{1}{n}\Sigma H_i$, $\bar{h} = \frac{1}{n}\Sigma h_i$, and $h_i = t_i y_i - \hat{E}(t_i\varepsilon_i)$.

This is exactly the normal equations system for the ALS estimator of Cheng and Schneeweiss (1996).

3 Case 2: No errors in the equation

Wolter and Fuller (1982) construct an estimator of the quadratic functional relationship when the whole of Ω is known. This corresponds to the case where there is no error in the equation but only measurement errors in the variables and the covariance matrix of the measurement errors is known. The estimator can be computed without any iterations. It can be generalized to the case of a polynomial functional relationship.

Let $z_i = (y_i, t_i')'$ and let W_i be the error covariance matrix of z_i and

$$\hat{W}_i = \begin{pmatrix} \sigma_\varepsilon^2 & \hat{v}_i' \\ \hat{v}_i & \hat{V}_i' \end{pmatrix}$$

its estimate. Furthermore let $M = \frac{1}{n}\ \Sigma\ z_i z_i'$, $m_{ty} = \frac{1}{n}\Sigma t_i y_i$, $M_{tt} = \frac{1}{n}\Sigma t_i t_i'$, $m_{yy} = \frac{1}{n}\Sigma y_i^2$, so that

$$M = \begin{pmatrix} m_{yy} & m_{ty}' \\ m_{ty} & M_{tt} \end{pmatrix}.$$

Then a generalization of Wolter and Fuller's estimator is given by

$$\hat{\beta} = (M_{tt} - \hat{\lambda}\bar{\hat{V}})^{-1}(m_{ty} - \hat{\lambda}\bar{\hat{v}}), \tag{5}$$

where $\hat{\lambda}$ is the smallest positve root (eigenvalue) of

$$det(M - \lambda \overline{\hat{W}}) = 0;$$

see also Moon and Gunst (1995) for the special case (N).

4 The asymptotic covariance matrix of $\hat{\beta}$ in case 2

Under general conditions $\sqrt{n}(\hat{\beta} - \beta)$ is asymptotically normally distributed with an asymptotic covariance matrix $\Sigma_{\hat{\beta}}$ which can be computed as follows.

First note that with $\theta = (1, -\beta')'$ and $\hat{\theta} = (1, -\hat{\beta}')'$ the estimating equation (5) for $\hat{\beta}$ can be written as

$$(M - \hat{\lambda}\overline{\hat{W}})\hat{\theta} = 0, \tag{6}$$

where $\hat{\lambda}$ is the smallest positive eigenvalue and $\hat{\theta}$ the corresponding eigenvector.

Let $\Delta M = M - EM$, where (see appendix)

$$EM = (\beta, I)'\overline{\zeta\zeta'}(\beta, I) + \bar{W}, \tag{7}$$

$\Delta\lambda = \hat{\lambda} - 1$, $\Delta\bar{W} = \overline{\hat{W}} - \bar{W}$, and $\Delta\hat{\theta} = \hat{\theta} - \theta$. For large n all these differences will be small in probability and the estimating equation (6) can be expanded as

$$(\Delta M - \Delta\lambda\bar{W} - \Delta\bar{W})\theta + (EM - \bar{W})\Delta\theta \approx 0.$$

This can be simplified with the help of (7) and using the fact that $(EM - \bar{W})\theta$ =0 to

$$(M - \overline{\hat{W}} - \Delta\lambda\bar{W})\theta \approx -(\beta, I)'\overline{\zeta\zeta'}(\beta, I)\Delta\theta.$$

Deleting the first equation of this system we get the following system for $\Delta\beta$:

$$\{(m_{ty}, M_{tt}) - (\overline{\hat{v}}, \overline{\hat{V}}) - (\hat{\lambda} - 1)(\bar{v}, \bar{V})\}\theta \approx \overline{\zeta\zeta'}\Delta\beta. \tag{8}$$

Now by (6)

$$\theta' M\hat{\theta} = \hat{\lambda}\theta'\overline{\hat{W}}\hat{\theta}$$

and again using $(EM - \bar{W})\theta = 0$ we similarly have

$$\theta' EM\hat{\theta} = \theta'\bar{W}\hat{\theta},$$

and taking differences

$$\theta'\Delta M\hat{\theta} = (\hat{\lambda} - 1)\theta'\overline{\hat{W}}\hat{\theta} + \theta'\overline{\hat{W}}\hat{\theta} - \theta'\bar{W}\hat{\theta}.$$

Thus

$$\Delta\lambda = \frac{\theta'(\Delta M - \Delta\bar{W})\hat{\theta}}{\theta'\overline{\hat{W}}\hat{\theta}} \approx \frac{\theta'(\Delta M - \Delta\bar{W})\theta}{\theta'\bar{W}\theta} = \frac{\theta'(M - \overline{\hat{W}})\theta}{\theta'\bar{W}\theta}.$$

Substituting this expression for $\hat{\lambda} - 1$ in (8) we get

$$\Delta\beta \approx (\overline{\zeta\zeta'})^{-1}\bar{\psi} \tag{9}$$

with $\bar{\psi} = \frac{1}{n}\Sigma\psi_i$ and

$$\begin{aligned}\psi_i &= \{(t_i y_i, t_i t_i') - (\hat{v}_i, \hat{V}_i) - \frac{\theta'(z_i z_i' - \hat{W}_i)\theta}{\theta'\bar{W}\theta}(\bar{v}, \bar{V})\}\theta \\ &= \{(0, I)(z_i z_i' - \hat{W}_i) - \frac{\theta'(z_i z_i' - \hat{W}_i)\theta}{\theta'\bar{W}\theta}(\bar{v}, \bar{V})\}\theta \qquad (10)\end{aligned}$$

Obviously $E\psi_i = 0$ (see appendix), and by the central limit theorem $\sqrt{n}\bar{\psi}$ converges in distribution to a normal distribution with covariance matrix

$$\lim E(\overline{\psi\psi'}).$$

Thus by (9) $\sqrt{n}(\hat{\beta} - \beta)$ also converges to a normal distribution with covariance matrix

$$\Sigma_{\hat{\beta}} = (\lim \overline{\zeta\zeta'})^{-1} \lim E(\overline{\psi\psi'})(\lim \overline{\zeta\zeta'})^{-1}$$

An estimate of the asymptotic covariance matrix of $\hat{\beta}$ is given by

$$\frac{1}{n}\hat{\Sigma}_{\hat{\beta}} = \frac{1}{n}\bar{H}^{-1}\overline{\hat{\psi}\hat{\psi}'}\bar{H}^{-1},$$

with

$$\hat{\psi}_i = t_i y_i - \hat{v}_i - (t_i t_i' - \hat{V}_i)\hat{\beta} - \frac{\hat{\theta}'(z_i z_i' - \hat{W}_i)\hat{\theta}}{\hat{\theta}'\overline{\hat{W}}\hat{\theta}}(\overline{\hat{v}} - \overline{\hat{V}}\hat{\beta}).$$

REFERENCES

CHAN, L. K. and MAK, M. K. (1985) On the polynomial functional relationship. *J. R. Statist. Soc. B*, **47**, 510-518.

CHENG, C. L. and SCHNEEWEISS, H. (1996) The polynomial regression with errors in the variables. *Disc. P.* 42 *of the Sonderforschungsbereich* 386, *University of Munich: Munich. To appear in JRSS*, *Ser. B.*

FULLER, W. A. (1987) Measurement Error Models. *New York: Wiley.*

GRILICHES, Z. and RINGSTAD, V. (1970) Errors-in-variables bias in nonlinear contexts. *Econometrica*, **38**, 368-370.

MOON, M.-S. and GUNST, R. (1995) Polynomial measurement error modelling. *Computational Statistics & Data Analysis*, **19**, 1-21.

WOLTER, K. H. and FULLER, W. A. (1982) Estimation of the quadratic errors-in-variables model. *Biometrika*, **69**, 175-182.

Appendix

We sketch a proof of (7) and $\mathsf{E}\psi_i = 0$.

First note that, by (2) and (4),

$$\begin{aligned} \mathsf{E} y_i^2 &= \beta' \zeta_i' \zeta_i \beta + \sigma_\varepsilon^2 \\ \mathsf{E}(t_i y_i) &= \mathsf{E} t_i \zeta_i' \beta + \mathsf{E}(e_i \varepsilon_i) = \zeta_i \zeta_i' \beta + v_i \\ \mathsf{E}(t_i t_i') &= \zeta_i \zeta_i' + V_i \end{aligned}$$

It follows that with $z_i = (y_i, t_i')'$

$$\mathsf{E}(z_i z_i') = \begin{pmatrix} \beta' \zeta_i \zeta_i' \beta & \beta' \zeta_i \zeta_i' \\ \zeta_i \zeta_i' \beta & \zeta_i \zeta_i' \end{pmatrix} + \begin{pmatrix} \sigma_\varepsilon^2 & v_i' \\ v_i & V_i \end{pmatrix}$$

$$\mathsf{E}(z_i z_i') = (\beta, I)' \zeta_i \zeta_i' (\beta, I) + W_i \tag{A1}$$

Averaging $(A1)$ over $i = 1, \ldots, n$ results in (7).

Now $(A1)$ also impies

$$\mathsf{E}(z_i z_i' - \hat{W}_i)\theta = (\beta, I)' \zeta_i \zeta_i' (\beta, I)(1, -\beta')' = 0.$$

It follows from (10) that $\mathsf{E}\psi_i = 0$.