ECONOMICS BULLETIN

Identification and Estimation of Structural-Change Models with Misclassification

Terence Tai-Leung Chong
The Chinese University of Hong Kong

Melvin Hinich
University of Texas at Austin

Kwan-To Wong
University of Hong Kong

Abstract

Consider a simple change-point model with a binary regressor. We examine the consistency of the change-point estimator when the regressor is subject to misclassification. It is found that the time of change can always be identified. Further, special cases where the structural parameters can also be identified are discussed. Simulation evidence is provided.

Citation: Chong, Terence Tai-Leung, Melvin Hinich, and Kwan-To Wong, (2007) "Identification and Estimation of Structural-Change Models with Misclassification." *Economics Bulletin*, Vol. 3, No. 36 pp. 1-19
Submitted: January 28, 2007. **Accepted:** August 28, 2007.
URL: http://economicsbulletin.vanderbilt.edu/2007/volume3/EB-07C20004A.pdf

1 Introduction

Measurement error is common in empirical data and can lead to serious estimation and inference problems. One of the earliest studies on this issue is Madansky (1959), who considers the problem of fitting a straight line when both variables are subject to errors. Levi (1973) shows that in a simple linear regression model without an intercept, if the explanatory variable is subject to errors, the regression estimates will be biased toward zero. Nelson (1995) obtains a similar attenuation bias in the multiple regression model. Chong and Lui (1998) show that the attenuation bias is nonlinear for fractionally integrated models.

The past couple of decades have witnessed a considerable effort to correct regression estimates for measurement error in the regressors. One of the commonly used methods to recover the true parameters is the instrumental-variable technique (Mahajan, 2006; Stefanski and Buzas, 1995; Fuller, 1987). Carroll et al. (2006) provide a comprehensive survey on the literature in nonlinear measurement error models. A special case of measurement error is misclassification, which occurs if the variable of interest is binary. Previous studies in misclassification include Küchenhoff et al. (2006) and Mahajan (2006), Dustmann and van-Soest (2001) and Poterba and Summers (1995, 1986).

In this paper, a structural-change model with a binary regressor measured with errors is examined. This model is different from conventional misclassification models in that it combines the problem of structural change[1] and misclassification, whereas conventional misclassification models do not consider changes in parameters. For example, an individual might develop an antibody at some point in time which causes him stop responding to a medicine[2]. Suppose for some reason that the respondent mis-report the information about treatment, or if there exist some systematic errors, then some observations will be contaminated. The point at issue is whether the time of shift in the response function can still be consistently estimated[3]. It will be shown in this paper that the change point can still be identified regardless of the existence of misclassification. Further, if the nature of the misclassification is known, then all the parameters can be recovered.

The remainder of this paper is organized as follows: Section 2 presents the model. Section 3 investigates the asymptotic properties of the least squares estimators for the change point and the pre- and post-shift parameters. Five special cases are discussed. Monte Carlo experiments are conducted in Section 4. Section 5 concludes the paper.

[1] There is a vast and growing literature on the structural-change models over the last decade. Recent studies by the author include Chong (2003, 2001).

[2] Similar examples include the development of the antibody of an insect to insecticide, or the mutation of an virus at some point in time.

[3] A related studies in this problem is Chang and Huang (1997).

2 The Model

Consider the model

$$\begin{aligned} y_t &= \alpha_1 (1 - x_t^*) + \gamma_1 x_t^* + u_t && \text{for } t \leq k_0, \\ y_t &= \alpha_2 (1 - x_t^*) + \gamma_2 x_t^* + u_t && \text{for } t > k_0, \end{aligned} \tag{1}$$

$t = 1, 2, ..., T$.

x_t^* is zero-one variable. For example, we let $x_t^* = 1$ if the respondent has taken the medicine at time t, and $x_t^* = 0$ otherwise.

(α_1, γ_1), (α_2, γ_2) are true structural parameters for $0 < t \leq k_0$ and $t > k_0$ respectively. In this paper, we exclude the case where $\alpha_1 = \alpha_2$ and $\gamma_1 = \gamma_2$. Let $k = [\tau T]$, where $[\cdot]$ is the greatest integer function, and $\tau \in [0, 1]$.

Suppose the true value of x_t^* is not perfectly measured and is approximated by an error-ridden measure called x_t. We define

$$p = \Pr(x_t = 0 | x_t^* = 1), \tag{2}$$

and

$$q = \Pr(x_t = 1 | x_t^* = 0). \tag{3}$$

The misclassification matrix (Küchenhoff et al, 2006) is therefore equal to

$$\Pi = \begin{pmatrix} 1 - q & p \\ q & 1 - p \end{pmatrix}. \tag{4}$$

Assume that:

$(A1)$ $\tau_0 = \dfrac{k_0}{T} \in K \subset (0, 1)$ where K is compact.

$(A2)$ $u_t \sim i.i.d.\left(0, \sigma_u^2\right)$, $\sigma_u^2 < \infty$.

$(A3)$ $x_t^* \sim i.i.d.$ which equals 1 with probability a and equals 0 with probability $1 - a$, where $0 \leq a \leq 1$.

$(A4)$ x_t^* are independent of u_t.

Assumption $(A1)$ states that the true change point belongs to a compact set in (0,1). This assumption is necessary because the least-squares estimators are not defined at the boundary of time domain. Assumptions $(A2) - (A4)$ describe the nature of the regressor and disturbance terms.

3 Estimation

3.1 Asymptotic Behavior of the Estimators

Model (1) can be rewritten as

$$
\begin{aligned}
y_t &= \alpha_1 + \beta_1 x_t^* + u_t \qquad \text{for } t \le k_0, \\
y_t &= \alpha_2 + \beta_2 x_t^* + u_t \qquad \text{for } t > k_0,
\end{aligned} \tag{5}
$$

where

$$\beta_1 = \gamma_1 - \alpha_1, \tag{6}$$

and

$$\beta_2 = \gamma_2 - \alpha_2. \tag{7}$$

Note that although the covariate is misclassified, the conditional mean $E(y_t|x_t)$ will still have a shift. As a result, we may still be able to estimate the true change point.

For any given $\tau \in (0,1)$, the pre-shift estimators are defined as

$$\widehat{\beta}_{1\tau} = \frac{\sum_{t=1}^{[\tau T]} (x_t - \overline{x})\, y_t}{\sum_{t=1}^{[\tau T]} (x_t - \overline{x})\, x_t}, \tag{8}$$

$$\widehat{\alpha}_{1\tau} = \frac{1}{[\tau T]} \left(\sum_{t=1}^{[\tau T]} y_t - \widehat{\beta}_{1\tau} \sum_{t=1}^{[\tau T]} x_t \right), \tag{9}$$

$$\widehat{\gamma}_{1\tau} = \widehat{\alpha}_{1\tau} + \widehat{\beta}_{1\tau}, \tag{10}$$

and the post-shift estimators are defined as

$$\widehat{\beta}_{2\tau} = \frac{\sum_{t=[\tau T]+1}^{T} (x_t - \overline{x})\, y_t}{\sum_{t=[\tau T]+1}^{T} (x_t - \overline{x})\, x_t}, \tag{11}$$

$$\widehat{\alpha}_{2\tau} = \frac{1}{T - [\tau T]} \left(\sum_{t=[\tau T]+1}^{T} y_t - \widehat{\beta}_{2\tau} \sum_{t=[\tau T]+1}^{T} x_t \right), \tag{12}$$

$$\widehat{\gamma}_{2\tau} = \widehat{\alpha}_{2\tau} + \widehat{\beta}_{2\tau}. \tag{13}$$

We define the change-point estimator as

$$\widehat{\tau} = Argmin_{\tau \in K} S_T(\tau), \tag{14}$$

where

$$S_T(\tau) = \sum_{t=1}^{[\tau T]} \left(y_t - \widehat{\alpha}_{1\tau} - \widehat{\beta}_{1\tau} x_t \right)^2 + \sum_{t=[\tau T]+1}^{T} \left(y_t - \widehat{\alpha}_{2\tau} - \widehat{\beta}_{2\tau} x_t \right)^2 \tag{15}$$

is the sum of squared residuals at τ.

The final pre- and post-shift estimators are evaluated at the change-point estimate $\widehat{\tau}$.

From the Appendix, $\frac{1}{T} S_T(\tau)$ converges uniformly to a non-stochastic function $h(\tau)$ such that $\sup_{\tau \in [0,1]} \left| \frac{1}{T} S_T(\tau) - h(\tau) \right| = o_p(1)$, where for $\tau \leq \tau_0$,

$$h(\tau) = h(\tau_0) + \Theta(\tau_0 - \tau) \frac{1 - \tau_0}{1 - \tau}, \tag{16}$$

and

$$\begin{aligned} h(\tau_0) &= \sigma_u^2 + a(1-a)\left(\tau_0 \beta_1^2 + (1-\tau_0)\beta_2^2\right) \\ &\times \left(1 - \frac{a(1-a)(1-p-q)^2}{(ap + (1-a)(1-q))(a(1-p) + (1-a)q)}\right), \end{aligned} \tag{17}$$

$$\begin{aligned} \Theta &= (\alpha_2 - \alpha_1 + a(\beta_2 - \beta_1))^2 \\ &+ \frac{(\beta_2 - \beta_1)^2 a^2 (1-a)^2 (1-p-q)^2}{(ap + (1-q)(1-a))(a(1-p) + q(1-a))}. \end{aligned}$$

Note that $h(\tau_0) \geq \sigma_u^2$. Thus, under fairly general conditions, the variance of the regression error u_t will be over-estimated.

The first and second derivatives of $h(\tau)$ are $h'(\tau) = -\Theta \frac{(1-\tau_0)^2}{(1-\tau)^2} \leq 0$ and $h''(\tau) = -2\Theta \frac{(1-\tau_0)^2}{(1-\tau)^3} \leq 0$ respectively. Hence, for $\tau \leq \tau_0$, $h(\tau)$ is non-increasing and concave. For $\tau > \tau_0, h(\tau) = h(\tau_0) + \Theta \frac{(\tau - \tau_0)\tau_0}{\tau}$, $h'(\tau) = \Theta \frac{\tau_0^2}{\tau^2} \geq 0$ and $h''(\tau) = -2\Theta \frac{\tau_0^2}{\tau^3} \leq 0$. Thus, for $\tau > \tau_0$, $h(\tau)$ is non-decreasing and concave. To summarize, for all $\tau \in [0,1]$, $\widehat{\alpha}_{1\tau}$, $\widehat{\alpha}_{2\tau}$, $\widehat{\gamma}_{1\tau}$ and $\widehat{\gamma}_{2\tau}$ are inconsistent estimates for α_1, α_2, γ_1 and γ_2 respectively. However, since the criterion function $\frac{1}{T} S_T(\tau)$ converges uniformly to a piecewise concave function $h(\tau)$ whose minimum takes place at the true change point, the change-point estimator is consistent.

Theorem 1: Under assumptions $(A1) - (A4)$, as $T \to \infty$, we have:

$$\widehat{\tau} \xrightarrow{p} \tau_0 \tag{18}$$

and

$$\widehat{\alpha}_{1\widehat{\tau}} \xrightarrow{p} \alpha_1 (1 - \lambda_b) + \gamma_1 \lambda_b, \tag{19}$$

$$\widehat{\gamma}_{1\widehat{\tau}} \xrightarrow{p} \alpha_1 (1 - \lambda_c) + \gamma_1 \lambda_c, \tag{20}$$

$$\widehat{\alpha}_{2\widehat{\tau}} \xrightarrow{p} \alpha_2 (1 - \lambda_b) + \gamma_2 \lambda_b, \tag{21}$$

$$\widehat{\gamma}_{2\widehat{\tau}} \xrightarrow{p} \alpha_2 (1 - \lambda_c) + \gamma_2 \lambda_c, \tag{22}$$

where

$$\lambda_a = \frac{a(1-a)(1-p-q)}{(ap + (1-a)(1-q))(a(1-p) + (1-a)q)}, \tag{23}$$

$$\lambda_b = \frac{ap}{ap + (1-a)(1-q)}, \tag{24}$$

$$\lambda_c = \lambda_a + \lambda_b = \frac{a(1-p)}{a(1-p) + (1-a)q}. \tag{25}$$

Proof. See Appendix.

Note that λ_a can be negative, whereas λ_b and λ_c are between zero and one. Theorem 1 states that the change point can be identified. However, the structural estimators converge to a convex combination of the regression coefficients. In general, $\widehat{\alpha}_{1\widehat{\tau}}$ and $\widehat{\alpha}_{2\widehat{\tau}}$ will be consistent if $\lambda_b = 0$, while $\widehat{\gamma}_{1\widehat{\tau}}$ and $\widehat{\gamma}_{2\widehat{\tau}}$ will be consistent if $q = 0$.

An inspection of Theorem 1 suggests that it is not possible to recover the true pre- and post-shift parameters without additional information. In our case, when λ_b and λ_c are known, the true pre- and post-shift parameters can be identified. Note that if p, q and a are known, then

$$\frac{\lambda_c \widehat{\alpha}_{1\widehat{\tau}} - \lambda_b \widehat{\gamma}_{1\widehat{\tau}}}{\lambda_a} \xrightarrow{p} \alpha_1, \tag{26}$$

$$\frac{\widehat{\gamma}_{1\widehat{\tau}} (1 - \lambda_b) - \widehat{\alpha}_{1\widehat{\tau}} (1 - \lambda_c)}{\lambda_a} \xrightarrow{p} \gamma_1, \tag{27}$$

$$\frac{\lambda_c \widehat{\alpha}_{2\widehat{\tau}} - \lambda_b \widehat{\gamma}_{2\widehat{\tau}}}{\lambda_a} \xrightarrow{p} \alpha_2, \tag{28}$$

$$\frac{\widehat{\gamma}_{2\widehat{\tau}} (1 - \lambda_b) - \widehat{\alpha}_{2\widehat{\tau}} (1 - \lambda_c)}{\lambda_a} \xrightarrow{p} \gamma_2. \tag{29}$$

In the cases where $\lambda_a = 0$, i.e., $a = 0$, $a = 1$, or $p + q = 1$, the coefficients cannot be recovered.

3.2 Special Cases

3.2.1 Case 1: $a = 0$ *or* $a = 1$

If there is only one category, two categories will be observed due to misclassification. The case where $a = 0$ is studied. When $a = 0$, we have $\lambda_a = \lambda_b = \lambda_c = 0$ and $\widehat{\alpha}_{1\widehat{\tau}} \xrightarrow{p} \alpha_1$, $\widehat{\gamma}_{1\widehat{\tau}} \xrightarrow{p} \alpha_1$, $\widehat{\alpha}_{2\widehat{\tau}} \xrightarrow{p} \alpha_2$, $\widehat{\gamma}_{2\widehat{\tau}} \xrightarrow{p} \alpha_2$.

The case where $a = 1$ has an opposite interpretation and is therefore skipped. In general, estimators for both categories converge to the true parameters of the existing category. The coefficients of the non-existing category cannot be identified even if the values of p and q are known.

3.2.2 Case 2: $p = 0$ *or* $q = 0$

Consider the case where $p = 0$, which implies that $\lambda_b = 0$ and

$$\lambda_a = \lambda_c = \frac{a}{a + (1 - a)\, q}.$$

In this case, we have

$$\widehat{\alpha}_{1\widehat{\tau}} \xrightarrow{p} \alpha_1,$$

$$\widehat{\gamma}_{1\widehat{\tau}} \xrightarrow{p} \alpha_1 (1 - \lambda_c) + \gamma_1 \lambda_c,$$

$$\widehat{\alpha}_{2\widehat{\tau}} \xrightarrow{p} \alpha_2,$$

$$\widehat{\gamma}_{2\widehat{\tau}} \xrightarrow{p} \alpha_2 (1 - \lambda_c) + \gamma_2 \lambda_c.$$

Thus, the parameters for one category can be identified. The structural estimators for another category converge to a convex combination of the regression coefficients. Further, if the values of q and a are known, then all the parameters can be identified. The case for $q = 0$ has an opposite interpretation and is therefore skipped.

3.2.3 Case 3: $p = 1$ *or* $q = 1$

When $p = 1$, i.e., the values in one category are always wrongly measured, we have $\lambda_c = 0$ and

$$\lambda_a = -\frac{a}{a + (1 - a)(1 - q)},$$

$$\lambda_b = -\lambda_a,$$

$$\widehat{\alpha}_{1\widehat{\tau}} \xrightarrow{p} \alpha_1 (1 - \lambda_b) + \gamma_1 \lambda_b,$$

$$\widehat{\gamma}_{1\widehat{\tau}} \xrightarrow{p} \alpha_1,$$

$$\widehat{\alpha}_{2\widehat{\tau}} \xrightarrow{p} \alpha_2 (1 - \lambda_b) + \gamma_2 \lambda_b,$$

$$\widehat{\gamma}_{2\widehat{\tau}} \xrightarrow{p} \alpha_2.$$

Note that the pre-shift and post-shift structural estimators are inconsistent. The estimators for the wrongly-measured category will converge to the true coefficient of another category. The estimate of the other category will be a convex combination of the true regression coefficients. If the information of p, q and a are available, then all the parameters can be retrieved. The case for $q = 1$ has an opposite interpretation and is therefore skipped.

3.2.4 Case 4: $p = q = 1$

When $p = q = 1$, the dichotomous covariate is always misclassified. In this case, $\lambda_a = -1$, $\lambda_b = 1$, $\lambda_c = 0$ and $\widehat{\alpha}_{1\widehat{\tau}} \xrightarrow{p} \gamma_1$, $\widehat{\gamma}_{1\widehat{\tau}} \xrightarrow{p} \alpha_1$, $\widehat{\alpha}_{2\widehat{\tau}} \xrightarrow{p} \gamma_2$, $\widehat{\gamma}_{2\widehat{\tau}} \xrightarrow{p} \alpha_2$. Thus, the estimator for one category will converge to the coefficient of another category.

3.2.5 Case 5: $p + q = 1$

When $p = 1 - q$, the two different categories are misclassified in such a way that the statistical properties of the two observed categories are identical. In this case, we have $\lambda_a = 0$, $\lambda_b = \lambda_c = a$ and

$$\widehat{\alpha}_{1\widehat{\tau}} \xrightarrow{p} \alpha_1 (1 - a) + \gamma_1 a,$$

$$\widehat{\gamma}_{1\widehat{\tau}} \xrightarrow{p} \alpha_1 (1 - a) + \gamma_1 a,$$

$$\widehat{\alpha}_{2\widehat{\tau}} \xrightarrow{p} \alpha_2 (1 - a) + \gamma_2 a,$$

$$\widehat{\gamma}_{2\widehat{\tau}} \xrightarrow{p} \alpha_2 (1 - a) + \gamma_2 a.$$

Thus, the two pre-shift estimators converge to the same point, and the two post-shift estimators also converge to the same value. Thus, even if the coefficients of the two categories are different, this difference cannot be observed due to the common statistical properties of the two observed categories. Further, since $\lambda_a = 0$, the true parameters cannot be identified.

4 Monte Carlo Experiments

This experiment verifies Theorem 1. Consider the model in Section 2:

$$\begin{aligned} y_t &= \alpha_1 (1 - x_t^*) + \gamma_1 x_t^* + u_t \qquad (t = 1, 2, ..., k_0), \\ y_t &= \alpha_2 (1 - x_t^*) + \gamma_2 x_t^* + u_t \qquad (t = k_0 + 1, k_0 + 2, ..., T). \end{aligned}$$

We perform the following experiment:
Let

$$T = 5000, \ k_0 = 2500, \ \tau_0 = \frac{k_0}{T} = .5.$$

$$u_t \sim Nid(0, 1),$$

$$x_t^* \sim i.i.d. \ Bernoulli((1, a), (0, 1 - a)),$$

$$x_t = x_t^* + \varepsilon_t.$$

$$\begin{aligned} \text{If } x_t^* &= 1, \text{ then } \varepsilon_t = -1 \text{ with probability } p \text{ and } \varepsilon_t = 0 \text{ with probability } (1 - p); \\ \text{if } x_t^* &= 0, \text{ then } \varepsilon_t = 1 \text{ with probability } q \text{ and } \varepsilon_t = 0 \text{ with probability } (1 - q). \end{aligned}$$

x_t^* and ε_t are independent of u_t. λ_a, λ_b and λ_c are defined as in Theorem 1.

For each value of a, p and q, we perform a single replication. The probability limits of the estimators are calculated under Theorem 1. The results of 20 cases are reported in Table 1.

Case 1 is the case without misclassification. Case 2 to case 9 are general cases. Case 10 to case 20 are special cases. Cases 10 and 11 correspond to the first special case. Note that estimators for both categories converge to the true parameters of the existing category. The coefficients of the non-existing category cannot be identified. Cases 12 and 13 correspond to special case 2. In case 12, when $q = 0$, γ_1 and γ_2 are identified. In case 13, when $p = 0$, α_1 and α_2 are identified. Cases 14 and 15 correspond to special case 3. Note that the pre-shift and post-shift structural estimators are inconsistent. The estimators for the wrongly-measured category converge to the true coefficient of another category. Case 6 is the fourth special case. The estimator for one category converges to the coefficient of another category. Cases 17 to 20 correspond to the last special case. The two pre-shift estimators converge to the same point, and the two post-shift estimators also converge to the same value. Note that the change point is consistently estimated in all cases. The simulated results in Table 1 largely conform to our theory.

Table 1: Performance of the estimators under various kinds of misclassifications

Case	1	2	3	4	5
(a,p,q)	(.5,0,0)	(.5,.3,.4)	(.5,.4,.3)	(.5,.2,.2)	(.3,.2,.2)
$(\alpha_1, \alpha_2, \gamma_1, \gamma_2)$	$(2, -1, 7, 9)$	$(2, -1, 7, 9)$	$(2, -1, 7, 9)$	$(10, 15, 10, 15)$	$(10, 15, 10, 15)$
λ_a	1	0.3030	0.3030	0.6	0.5348
λ_b	0	0.3333	0.3636	0.2	0.0968
λ_c	1	0.6363	0.6667	0.8	0.6316
plim $\widehat{\alpha}_{1\widehat{\tau}}$	2	3.6667	3.8090	10	10
plim $\widehat{\alpha}_{2\widehat{\tau}}$	-1	2.3333	2.7796	15	15
plim $\widehat{\gamma}_{1\widehat{\tau}}$	7	5.1818	5.3333	10	10
plim $\widehat{\gamma}_{2\widehat{\tau}}$	9	5.3995	5.6667	15	15
$\widehat{\alpha}_{1\widehat{\tau}}$	2.0588	3.7218	3.8182	9.9978	9.9590
$\widehat{\alpha}_{2\widehat{\tau}}$	-0.9690	2.1099	2.6364	15.01	15.00
$\widehat{\gamma}_{1\widehat{\tau}}$	6.9732	5.1542	5.3769	9.9945	10.02
$\widehat{\gamma}_{2\widehat{\tau}}$	8.9776	5.1455	5.7299	15.01	15.03
$\widehat{\tau}$	0.5000	0.5060	0.5000	0.5000	0.5002

Case	6	7	8	9	10
(a,p,q)	(.5,.7,.4)	(.5,.6,.6)	(.3,.2,.4)	(.3,.4,.2)	(0,.2,.2)
$(\alpha_1, \alpha_2, \gamma_1, \gamma_2)$	$(2, -1, 7, 9)$	$(2, -1, 7, 9)$	$(2, -1, 7, 9)$	$(2, -1, 7, 9)$	$(10, 15, 15, 23)$
λ_a	-0.1099	-0.2	0.3366	0.3860	0
λ_b	0.5385	0.6	0.125	0.1765	0
λ_c	0.4286	0.4	0.4615	0.5625	0
plim $\widehat{\alpha}_{1\widehat{\tau}}$	4.6923	5	2.625	2.8824	10
plim $\widehat{\alpha}_{2\widehat{\tau}}$	4.3846	5	0.25	0.7647	15
plim $\widehat{\gamma}_{1\widehat{\tau}}$	4.1429	4	4.3077	4.8125	10
plim $\widehat{\gamma}_{2\widehat{\tau}}$	3.2857	3	3.6154	4.625	15
$\widehat{\alpha}_{1\widehat{\tau}}$	4.8293	4.8857	2.6294	2.8476	10.01
$\widehat{\alpha}_{2\widehat{\tau}}$	4.5149	4.9573	0.2770	0.7003	14.96
$\widehat{\gamma}_{1\widehat{\tau}}$	4.0639	3.9620	4.4003	4.8606	10.04
$\widehat{\gamma}_{2\widehat{\tau}}$	3.0039	3.0244	3.6584	4.3695	14.99
$\widehat{\tau}$	0.4998	0.4960	0.4970	0.5076	0.5000

Table 1 cont.

Case	11	12	13	14	15
(a,p,q)	(1,.2,.2)	(.5,.3,0)	(.5,0,.4)	(.5,1,.3)	(.5,.3,1)
$(\alpha_1, \alpha_2, \gamma_1, \gamma_2)$	$(10, 15, 15, 23)$	$(2, -1, 7, 9)$	$(2, -1, 7, 9)$	$(10, 15, 15, 23)$	$(5, 8, 15, 23)$
λ_a	0	0.625	0.7143	-0.5882	-0.5882
λ_b	1	0.375	0	0.5882	1
λ_c	1	1	0.7143	0	0.4118
plim $\widehat{\alpha}_{1\widehat{\tau}}$	15	3.875	2	12.94	15
plim $\widehat{\alpha}_{2\widehat{\tau}}$	23	2.75	-1	19.76	23
plim $\widehat{\gamma}_{1\widehat{\tau}}$	15	7	5.5714	10	12.06
plim $\widehat{\gamma}_{2\widehat{\tau}}$	23	9	6.1429	15	18.29
$\widehat{\alpha}_{1\widehat{\tau}}$	14.89	3.8664	2.0139	12.89	14.99
$\widehat{\alpha}_{2\widehat{\tau}}$	22.96	2.7719	-1.0173	19.54	23.01
$\widehat{\gamma}_{1\widehat{\tau}}$	14.99	6.9242	5.5294	10.01	11.99
$\widehat{\gamma}_{2\widehat{\tau}}$	23.00	8.9698	6.0438	14.99	18.15
$\widehat{\tau}$	0.5000	0.5000	0.5000	0.5002	0.5000

Case	16	17	18	19	20
(a,p,q)	(.5,1,1)	(.5,.6,.4)	(.5,.6,.4)	(.5,.5,.5)	(.5,.3,.7)
$(\alpha_1, \alpha_2, \gamma_1, \gamma_2)$	$(2, -1, 7, 9)$	$(10, 15, 15, 23)$	$(2, -1, 7, 9)$	$(2, -1, 7, 9)$	$(2, -1, 7, 9)$
λ_a	-1	0	0	0	0
λ_b	1	0.5	0.5	0.5	0.5
λ_c	0	0.5	0.5	0.5	0.5
plim $\widehat{\alpha}_{1\widehat{\tau}}$	7	12.5	4.5	4.5	4.5
plim $\widehat{\alpha}_{2\widehat{\tau}}$	9	19	4	4	4
plim $\widehat{\gamma}_{1\widehat{\tau}}$	2	12.5	4.5	4.5	4.5
plim $\widehat{\gamma}_{2\widehat{\tau}}$	-1	19	4	4	4
$\widehat{\alpha}_{1\widehat{\tau}}$	6.993	12.47	4.4957	4.5847	4.3995
$\widehat{\alpha}_{2\widehat{\tau}}$	9.019	19.04	3.9884	3.7918	3.5123
$\widehat{\gamma}_{1\widehat{\tau}}$	1.980	12.35	4.3579	4.5788	4.4196
$\widehat{\gamma}_{2\widehat{\tau}}$	-0.99	19.30	3.7532	4.1638	4.0422
$\widehat{\tau}$	0.5000	0.5000	0.5110	0.4960	0.5070

5 Conclusion

In sum, this paper studies a structural-change model with the regressor being a zero-one variable subject to misclassification. This kind of model is new, in the sense that none of the previous studies in misclassification has considered the structural-change problem. The interest of this paper lies primarily in the unknown change point. Despite the fact that the data are contaminated, and the existence of a non-zero correlation between the latent variable, x_t^* and latent random error ε_t, it is shown that the time of change can still be identified. Further, it is also shown that the true structural parameters can be extracted from the information of p, q and a. Special cases of our model, as well as Monte Carlo evidence are provided to help illustrating the generic identifiability of the change point in the presence of classification errors. Our results are in line with Chong (2003), who has shown that the consistency of the change-point estimator is preserved when the regression model is misspecified.

APPENDIX

Proof of Theorem 1:
Let

$$\varepsilon_t = x_t - x_t^*.$$

Note that

$$E(x_t^{*2}) = E(x_t^*) = 1 \times \Pr(x_t^* = 1) + 0 \times \Pr(x_t^* = 0) = a.$$

$$Var(x_t^*) = a(1-a).$$

$$\begin{aligned} E(\varepsilon_t) &= E(\varepsilon_t | x_t^* = 1)\Pr(x_t^* = 1) + E(\varepsilon_t | x_t^* = 0)\Pr(x_t^* = 0) \\ &= [p \times (-1) + (1-p) \times 0]\, a + [q \times (1) + (1-q) \times 0]\,(1-a) \\ &= -ap + (1-a)q. \end{aligned}$$

$$\begin{aligned} E(\varepsilon_t^2) &= E\left(\varepsilon_t^2 | x_t^* = 1\right)\Pr(x_t^* = 1) + E\left(\varepsilon_t^2 | x_t^* = 0\right)\Pr(x_t^* = 0) \\ &= \left[p \times (-1)^2 + (1-p) \times 0^2\right] a + \left[q \times (1)^2 + (1-q) \times 0^2\right](1-a) \\ &= ap + (1-a)q. \end{aligned}$$

$$\begin{aligned} Var(\varepsilon_t) &= E(\varepsilon_t^2) - E^2(\varepsilon_t) = ap + (1-a)q - (-ap + (1-a)q)^2 \\ &= 2ap - (ap - q(1-a))(a(p+q) + 1 - q). \end{aligned}$$

$$\begin{aligned} E(x_t^*\varepsilon_t) &= E\left(x_t^*\varepsilon_t|x_t^*=1\right)\Pr\left(x_t^*=1\right)+E\left(x_t^*\varepsilon_t|x_t^*=0\right)\Pr\left(x_t^*=0\right) \\ &= E\left(\varepsilon_t|x_t^*=1\right)\Pr\left(x_t^*=1\right) \\ &= \left[p\times(-1)+(1-p)\times 0\right]a \\ &= -ap. \end{aligned}$$

$$Cov\left(x_t^*,\varepsilon_t\right)=E(x_t^*\varepsilon_t)-E(x_t^*)E(\varepsilon_t)=-ap-a\left(-ap+(1-a)q\right)=-a\left(p+q\right)\left(1-a\right).$$

$$E(x_t)=E(x_t^*)+E(\varepsilon_t)=a\left(1-p\right)+(1-a)q.$$

$$E(x_t^2)=E(x_t^{*2})+E(\varepsilon_t^2)+2E(x_t^*\varepsilon_t)=a+ap+(1-a)q-2ap=a\left(1-p\right)+\left(1-a\right)q.$$

$$E\left(x_t\varepsilon_t\right)=E(x_t^*\varepsilon_t)+E(\varepsilon_t^2)=-ap+ap+(1-a)q=(1-a)q.$$

$$\begin{aligned} Var\left(x_t\right) &= a\left(1-a\right)+2ap-\left(ap-q+qa\right)\left(ap+1-q+qa\right)-2a\left(p+q\right)\left(1-a\right) \\ &= \left(ap+\left(1-a\right)\left(1-q\right)\right)\left(a\left(1-p\right)+\left(1-a\right)q\right). \end{aligned}$$

For $\tau\in[0,1]$,

$$S_{**}\left(\tau\right)\overset{def}{=}\frac{1}{T}\sum_{t=1}^{[\tau T]}\left(x_t^*-\overline{x^*}\right)^2\overset{p}{\to}\tau Var(x_t^*),$$

$$S_{\varepsilon\varepsilon}\left(\tau\right)\overset{def}{=}\frac{1}{T}\sum_{t=1}^{[\tau T]}\left(\varepsilon_t-\overline{\varepsilon}\right)^2\overset{p}{\to}\tau Var(\varepsilon_t)$$

uniformly.

These results bound the variation of the stochastic insignificant terms and will be utilized in the proof of the uniform convergence result below.

Let

$$\lambda_a=\frac{a\left(1-a\right)\left(1-p-q\right)}{\left(ap+\left(1-a\right)\left(1-q\right)\right)\left(a\left(1-p\right)+\left(1-a\right)q\right)},$$

$$\lambda_b=\frac{ap}{ap+\left(1-a\right)\left(1-q\right)},$$

$$\lambda_c=\lambda_a+\lambda_b=\frac{a\left(1-p\right)}{a\left(1-p\right)+\left(1-a\right)q},$$

$$\Gamma_1(\tau) = \beta_1 \frac{\tau_0 - \tau}{1-\tau} + \beta_2 \frac{1-\tau_0}{1-\tau},$$

$$\Psi_1(\tau) = \alpha_1 \frac{\tau_0 - \tau}{1-\tau} + \alpha_2 \frac{1-\tau_0}{1-\tau},$$

$$\Lambda_1(\tau) = \gamma_1 \frac{\tau_0 - \tau}{1-\tau} + \gamma_2 \frac{1-\tau_0}{1-\tau},$$

$$\Gamma_2(\tau) = \beta_1 \frac{\tau_0}{\tau} + \beta_2 \frac{\tau - \tau_0}{\tau},$$

$$\Psi_2(\tau) = \alpha_1 \frac{\tau_0}{\tau} + \alpha_2 \frac{\tau - \tau_0}{\tau},$$

$$\Lambda_2(\tau) = \gamma_1 \frac{\tau_0}{\tau} + \gamma_2 \frac{\tau - \tau_0}{\tau}.$$

For $\tau \le \tau_0$,

$$\widehat{\beta}_{1\tau} = \frac{\sum_{t=1}^{[\tau T]} (x_t - \overline{x})(\alpha_1 + \beta_1 x_t^* + u_t)}{\sum_{t=1}^{[\tau T]} (x_t - \overline{x}) x_t} \xrightarrow{p} \beta_1 \lambda_a.$$

$$\begin{aligned}
\widehat{\alpha}_{1\tau} &= \frac{1}{[\tau T]} \left(\sum_{t=1}^{[\tau T]} (\alpha_1 + \beta_1 x_t^* + u_t) - \widehat{\beta}_{1\tau} \sum_{t=1}^{[\tau T]} (x_t^* + \varepsilon_t) \right) \\
&\xrightarrow{p} \alpha_1 + \beta_1 \left[a(1-\lambda_a) + (ap - (1-a)q)\lambda_a \right] \\
&= \alpha_1 (1-\lambda_b) + \gamma_1 \lambda_b.
\end{aligned}$$

$$\widehat{\beta}_{2\tau} \xrightarrow{p} \left(\beta_1 \frac{\tau_0 - \tau}{1-\tau} + \beta_2 \frac{1-\tau_0}{1-\tau} \right) \lambda_a = \Gamma_1(\tau) \lambda_a.$$

$$\begin{aligned}
\widehat{\alpha}_{2\tau} &= \frac{1}{T - [\tau T]} \left(\sum_{t=[\tau T]+1}^{[\tau_0 T]} (\alpha_1 + \beta_1 x_t^* + u_t) + \sum_{t=[\tau_0 T]+1}^{T} (\alpha_2 + \beta_2 x_t^* + u_t) - \widehat{\beta}_{2\tau} \sum_{t=[\tau T]+1}^{T} (x_t^* + \varepsilon_t) \right) \\
&\xrightarrow{p} \Psi_1(\tau) + \Gamma_1(\tau) \left[(1-\lambda_a) E(x_t^*) - \lambda_a E(\varepsilon_t) \right] \\
&= \Psi_1(\tau) + \Gamma_1(\tau) \lambda_b \\
&= \Psi_1(\tau)(1-\lambda_b) + \Lambda_1(\tau) \lambda_b.
\end{aligned}$$

$$\tfrac{1}{T} S_T(\tau) = \tfrac{1}{T} \sum_{t=1}^{[\tau T]} \left(y_t - \widehat{\alpha}_{1\tau} - \widehat{\beta}_{1\tau} x_t \right)^2 + \tfrac{1}{T} \sum_{t=[\tau T]+1}^{k_0} \left(y_t - \widehat{\alpha}_{2\tau} - \widehat{\beta}_{2\tau} x_t \right)^2 + \tfrac{1}{T} \sum_{t=k_0+1}^{T} \left(y_t - \widehat{\alpha}_{2\tau} - \widehat{\beta}_{2\tau} x_t \right)^2$$

$$= \frac{1}{T}\sum_{t=1}^{[\tau T]}\left((\alpha_1-\widehat{\alpha}_{1\tau})+\beta_1 x_t^*+u_t-\widehat{\beta}_{1\tau}x_t\right)^2+\frac{1}{T}\sum_{t=[\tau T]+1}^{k_0}\left((\alpha_1-\widehat{\alpha}_{2\tau})+\beta_1 x_t^*+u_t-\widehat{\beta}_{2\tau}x_t\right)^2$$
$$+\frac{1}{T}\sum_{t=k_0+1}^{T}\left((\alpha_2-\widehat{\alpha}_{2\tau})+\beta_2 x_t^*+u_t-\widehat{\beta}_{2\tau}x_t\right)^2$$

$$= \frac{1}{T}\sum_{t=1}^{T}u_t^2+\frac{1}{T}\sum_{t=1}^{[\tau T]}\left(\beta_1 x_t^*-\beta_1\left(\lambda_b+\lambda_a x_t\right)\right)^2$$
$$+\frac{1}{T}\sum_{t=[\tau T]+1}^{k_0}\left(\left(\alpha_1-\alpha_2\right)\frac{1-\tau_0}{1-\tau}+\beta_1 x_t^*-\Gamma_1\left(\tau\right)\left(\lambda_b+\lambda_a x_t\right)\right)^2$$
$$+\frac{1}{T}\sum_{t=k_0+1}^{T}\left(\left(\alpha_2-\alpha_1\right)\frac{\tau_0-\tau}{1-\tau}+\beta_2 x_t^*-\Gamma_1\left(\tau\right)\left(\lambda_b+\lambda_a x_t\right)\right)^2+o_p\left(1\right)$$

$$\xrightarrow{p}\sigma_u^2+\tau\beta_1^2 E\left(x_t^*-\left(\lambda_b+\lambda_a x_t\right)\right)^2+\left(\tau_0-\tau\right)E\left(\left(\alpha_1-\alpha_2\right)\frac{1-\tau_0}{1-\tau}+\beta_1 x_t^*-\Gamma_1\left(\tau\right)\left(\lambda_b+\lambda_a x_t\right)\right)^2$$
$$+\left(1-\tau_0\right)E\left(\left(\alpha_2-\alpha_1\right)\frac{\tau_0-\tau}{1-\tau}+\beta_2 x_t^*-\Gamma_1\left(\tau\right)\left(\lambda_b+\lambda_a x_t\right)\right)^2$$

$$=\sigma_u^2+\left(\tau_0\beta_1^2+\left(1-\tau_0\right)\beta_2^2\right)E\left(x_t^{*2}\right)+\left(\tau_0-\tau\right)\frac{1-\tau_0}{1-\tau}\left(\left(\alpha_2-\alpha_1\right)+\left(\beta_2-\beta_1\right)E\left(x_t^*\right)\right)^2$$
$$+\left(\tau\beta_1^2+\left(1-\tau\right)\Gamma_1^2\left(\tau\right)\right)E\left(\lambda_b+\lambda_a x_t\right)^2-\left(\tau_0-\tau\right)\frac{1-\tau_0}{1-\tau}\left(\beta_2-\beta_1\right)^2 E^2\left(x_t^*\right)$$
$$-2\left[\tau\beta_1^2+\left(1-\tau\right)\Gamma_1^2\left(\tau\right)\right]E\left(x_t^*\left(\lambda_b+\lambda_a x_t\right)\right)$$

$$=\sigma_u^2+\left(\tau_0-\tau\right)\frac{1-\tau_0}{1-\tau}\left(\left(\alpha_2-\alpha_1\right)+\left(\beta_2-\beta_1\right)a\right)^2$$
$$+\left(\tau_0\beta_1^2+\left(1-\tau_0\right)\beta_2^2\right)a-\left(\tau_0-\tau\right)\frac{1-\tau_0}{1-\tau}\left(\beta_2-\beta_1\right)^2 a^2-\left(\tau\beta_1^2+\left(1-\tau\right)\Gamma_1^2\left(\tau\right)\right)a^2$$
$$+\left(\tau\beta_1^2+\left(1-\tau\right)\Gamma_1^2\left(\tau\right)\right)\left[E\left(\lambda_b+\lambda_a x_t\right)^2-2E\left(x_t^*\left(\lambda_b+\lambda_a x_t\right)\right)+a^2\right]$$

$$=\sigma_u^2+\left(\tau_0-\tau\right)\frac{1-\tau_0}{1-\tau}\left(\left(\alpha_2-\alpha_1\right)+\left(\beta_2-\beta_1\right)a\right)^2+\left(\tau_0\beta_1^2+\left(1-\tau_0\right)\beta_2^2\right)\left(1-a\right)a$$
$$-\frac{a^2\left(1-a\right)^2\left(1-p-q\right)^2}{\left(ap+\left(1-a\right)\left(1-q\right)\right)\left(a\left(1-p\right)+\left(1-a\right)q\right)}\left(\tau\beta_1^2+\left(1-\tau\right)\Gamma_1^2\left(\tau\right)\right)$$

$$\stackrel{def}{=} h\left(\tau\right).$$

Note that

$$h'\left(\tau\right)=-\Theta\frac{\left(1-\tau_0\right)^2}{\left(1-\tau\right)^2}\leq 0,$$

$$h''\left(\tau\right)=-2\Theta\frac{\left(1-\tau_0\right)^2}{\left(1-\tau\right)^3}\leq 0,$$

where

$$\Theta = (\alpha_2 - \alpha_1 + a(\beta_2 - \beta_1))^2 + \frac{(\beta_2 - \beta_1)^2 a^2 (1-a)^2 (1-p-q)^2}{(ap + (1-q)(1-a))(a(1-p) + q(1-a))}.$$

For $\tau > \tau_0$,

$$\widehat{\beta}_{1\tau} \xrightarrow{p} \Gamma_2(\tau)\lambda_a.$$

$$\begin{aligned}
\widehat{\alpha}_{1\tau} &= \frac{1}{[\tau T]}\left(\sum_{t=1}^{[\tau_0 T]} y_t - \widehat{\beta}_{1\tau}\sum_{t=1}^{[\tau_0 T]} x_t\right) + \frac{1}{[\tau T]}\left(\sum_{t=[\tau_0 T]+1}^{[\tau T]} y_t - \widehat{\beta}_{1\tau}\sum_{t=[\tau_0 T]+1}^{[\tau T]} x_t\right) \\
&\xrightarrow{p} \frac{\tau_0}{\tau}\left(\alpha_1 + \beta_1 E(x_t^*) + E(u_t) - \Gamma_2(\tau)\lambda_a(E(x_t^*) + E(\varepsilon_t))\right) \\
&\quad + \frac{\tau - \tau_0}{\tau}\left(\alpha_2 + \beta_2 E(x_t^*) + E(u_t) - \Gamma_2(\tau)\lambda_a(E(x_t^*) + E(\varepsilon_t))\right) \\
&= \Psi_2(\tau) + [\Gamma_2(\tau) - \Gamma_2(\tau)\lambda_a]\,a - (-ap + (1-a)q)\,\Gamma_2(\tau)\lambda_a \\
&= \Psi_2(\tau)(1-\lambda_b) + \Lambda_2(\tau)\lambda_b.
\end{aligned}$$

$$\widehat{\gamma}_{1\tau} = \widehat{\alpha}_{1\tau} + \widehat{\beta}_{1\tau} \xrightarrow{p} \Psi_2(\tau)(1-\lambda_c) + \Lambda_2(\tau)\lambda_c.$$

$$\widehat{\beta}_{2\tau} \xrightarrow{p} \beta_2\lambda_a.$$

$$\begin{aligned}
\widehat{\alpha}_{2\tau} &= \frac{1}{T - [\tau T]}\left(\sum_{t=[\tau T]+1}^{T}(\alpha_2 + \beta_2 x_t^* + u_t) - \widehat{\beta}_{2\tau}\sum_{t=[\tau T]+1}^{T}(x_t^* + \varepsilon_t)\right) \\
&= \alpha_2 + \frac{1}{T - [\tau T]}\left(\beta_2\sum_{t=[\tau T]+1}^{T} x_t^* + \sum_{t=[\tau T]+1}^{T} u_t - \widehat{\beta}_{2\tau}\sum_{t=[\tau T]+1}^{T}(x_t^* + \varepsilon_t)\right) \\
&\xrightarrow{p} \alpha_2 + \beta_2\left(a - \frac{a(1-a)(1-p-q)}{(ap + (1-a)(1-q))(a(1-p) + (1-a)q)}(a(1-p) + (1-a)q)\right) \\
&= \alpha_2(1-\lambda_b) + \gamma_2\lambda_b.
\end{aligned}$$

$$\widehat{\gamma}_{2\tau} = \widehat{\alpha}_{2\tau} + \widehat{\beta}_{2\tau} \xrightarrow{p} \alpha_2(1-\lambda_c) + \gamma_2\lambda_c.$$

$$\begin{aligned}
&\tfrac{1}{T} S_T(\tau) \\
&\quad = \tfrac{1}{T}\sum_{t=1}^{[\tau_0 T]}\left((\alpha_1 - \widehat{\alpha}_{1\tau}) + \beta_1 x_t^* + u_t - \widehat{\beta}_{1\tau}x_t\right)^2 + \tfrac{1}{T}\sum_{t=k_0+1}^{[\tau T]}\left((\alpha_2 - \widehat{\alpha}_{1\tau}) + \beta_2 x_t^* + u_t - \widehat{\beta}_{1\tau}x_t\right)^2
\end{aligned}$$

$$+\frac{1}{T}\sum_{t=[\tau T]+1}^{T}\left((\alpha_2-\widehat{\alpha}_{2\tau})+\beta_2 x_t^*+u_t-\widehat{\beta}_{2\tau}x_t\right)^2$$

$$\xrightarrow{p}\sigma_u^2+(1-\tau)\beta_2^2E(x_t^*-(\lambda_b+\lambda_a x_t))^2+(\tau-\tau_0)E\left((\alpha_2-\alpha_1)\frac{\tau_0}{\tau}+\beta_2 x_t^*-\Gamma_2(\tau)(\lambda_b+\lambda_a x_t)\right)^2$$
$$+\tau_0 E\left((\alpha_1-\alpha_2)\frac{\tau-\tau_0}{\tau}+\beta_1 x_t^*-\Gamma_2(\tau)(\lambda_b+\lambda_a x_t)\right)^2$$

$$=\sigma_u^2+\frac{(\tau-\tau_0)\tau_0}{\tau}(\alpha_2-\alpha_1+a(\beta_2-\beta_1))^2$$
$$+\left((1-\tau_0)\beta_2^2+\tau_0\beta_1^2\right)a-\frac{(\tau-\tau_0)\tau_0}{\tau}(\beta_2-\beta_1)^2(E(x_t^*))^2-\left((1-\tau)\beta_2^2+\tau\Gamma_2^2(\tau)\right)a^2$$
$$\left((1-\tau)\beta_2^2+\tau\Gamma_2^2(\tau)\right)\left[E(\lambda_b+\lambda_a x_t)^2-2E(x_t^*(\lambda_b+\lambda_a x_t))+a^2\right]$$

$$=\sigma_u^2+\frac{(\tau-\tau_0)\tau_0}{\tau}(\alpha_2-\alpha_1+a(\beta_2-\beta_1))^2+\left(\tau_0\beta_1^2+(1-\tau_0)\beta_2^2\right)(1-a)a$$
$$-\frac{a^2(1-a)^2(1-p-q)^2}{(ap+(1-a)(1-q))(a(1-p)+(1-a)q)}\left((1-\tau)\beta_2^2+\tau\Gamma_2^2(\tau)\right)$$

$$\stackrel{def}{=}h(\tau).$$

$$h'(\tau)=\Theta\frac{\tau_0^2}{\tau^2}\geq 0.$$

$$h''(\tau)=-2\Theta\frac{\tau_0^2}{\tau^3}\leq 0.$$

Thus, $\frac{1}{T}S_T(\tau)$ converges pointwise to a piecewise concave function $h(\tau)$, with the unique minimum at the true change point. Under assumptions $(A1)$ to $(A4)$, the uniform convergence result follows, i.e.,

$$\sup_{\tau\in[0,1]}\left|\frac{1}{T}S_T(\tau)-h(\tau)\right|=o_p(1).$$

The consistency results of Chong (2001) applies. The change-point estimator is T-consistent, and the pre- and post-shift estimators are $\sqrt{T}$ -consistent. Q.E.D.

REFERENCES

1. Carroll, R., D. Ruppert and L. Stefanski and C. M. Crainiceanu (2006), *Measurement Error in Non-linear Models. A Modern Perspective*, Second Edition. London: Chapman and Hall.

2. Chang, Y. P. and W. T. Huang (1997), "Inferences for the Linear Errors-in-Variables with Changepoint Models," *Journal of the American Statistical Association* 92 , 171–178.

3. Chong, T. T. L. (2003), "Generic Consistency of the Break-Point Estimator Under Specification Errors," *Econometrics Journal* 6, 167-192.

4. Chong, T. T. L. (2001), "Structural Change in AR(1) Models," *Econometric Theory* 17, 87-155.

5. Chong, T. T .L. and G.C.S. Lui (1999), "Estimating the Fractionally Integrated Process in the Presence of Measurement Errors" *Economics Letters* 63, 285-294.

6. Dustmann, C and A. van-Soest (2001), "Language Fluency and Earnings: Estimation with Misclassified Language Indicators," *Review of Economics and Statistics*, 83(4), 663-674.

7. Fuller, W. A. (1987), *Measurement Error Models*, New York: John Wiley.

8. Küchenhoff, H., S. M. Mwalili, and E. Lesaffre (2006), "A General Method for Dealing with Misclassification in Regression: The Misclassification SIMEX," *Biometrics* 62, 85–96.

9. Levi, M. D. (1973), "Errors in the Variables bias in the Presence of Correctly Measured Variables," *Econometrica* 41, 985-986.

10. Madansky, A. (1959), "The Fitting of Straight Lines with Both Variables are subject to Error," *Journal of the American Statistical Association* 54, 173-205.

11. Mahajan, A. (2006), "Identification and Estimation of Regression Models with Misclassification," *Econometrica* 74, 631–665.

12. Nelson, D. B. (1995), "Vector Attenuation Bias in the Classical Errors-in-Variables Model," *Economics Letters* 49, 345-349.

13. Poterba, J. M. and L. H. Summers (1986), "Reporting Errors and Labor Market Dynamics," *Econometrica* 54, 1319-1338.

14. Poterba, J. M. and L. H. Summers (1995), "Unemployment Benefits and Labor Market Transitions: A Multinomial Logit Model With Errors in Classification," *Review of Economics and Statistics* 77, 207-216.

15. Stefanski, L. A.and J. S. Buzas (1995), "Instrumental Variable Estimation in Binary Regression Measurement Error Models," *Journal of the American Statistical Association* 90, 541–550.