IMS Collections
Nonparametrics and Robustness in Modern Statistical Inference and Time Series Analysis: A Festschrift in honor of Professor Jana Jurečková
Vol. 7 (2010) 224–234

DOI: 10.1214/10-IMSCOLL722

The Theil–Sen estimator in a measurement error perspective*

Pranab K. Sen[1] **and A. K. Md. Ehsanes Saleh**[2]

University of North Carolina at Chapel Hill and Carleton University, Ottawa

Abstract: In a simple measurement error regression model, the classical least squares estimator of the slope parameter consistently estimates a discounted slope, though sans normality, some other properties may not hold. It is shown that for a broader class of error distributions, the Theil–Sen estimator, albeit nonlinear, is a median-unbiased, consistent and robust estimator of the same discounted parameter. For a general class of nonlinear (including $R-$, $M-$ and $L-$ estimators), study of asymptotic properties is greatly facilitated by using some uniform asymptotic linearity results, which are, in turn, based on contiguity of probability measures. This contiguity is established in a measurement error model under broader distributional assumptions. Some asymptotic properties of the Theil–Sen estimator are studied under slightly different regularity conditions in a direct way bypassing the contiguity approach.

1. Introduction

For the simple regression model $Y = \theta + \beta x + e$, with nonstochastic regressors, the estimator of the slope parameter β based on the Kendall tau statistic, known as the Theil–Sen estimator (TSE), is robust, median-unbiased and it provides a distribution-free confidence interval for β (Sen [11]). When the regressors are themselves stochastic and, in addition, they are subject to measurement errors (ME), like the classical least squares estimator (LSE), the TSE does not estimate the slope unbiasedly or even consistently. The LSE, under some additional regularity assumptions, estimate a discounted regression parameter $\gamma = \kappa\beta$, where the discounting factor κ is the variance ratio of the unobserved and observed regressors. In this ME setup, there are some basic qualms:

(i) Sans the normality of the errors, when does the LSE in a ME setup estimate γ consistently and median-unbiasedly?
(ii) In the same ME setup, when does the TSE estimate γ consistently and median-unbiasedly?

Researchers in the past have relied heavily on normality of the errors including the ME component (Fuller [1]), albeit in real applications such stringent assumptions

*Supported by Boshamer Research Fund at the University of North Carolina at Chapel Hill, and Discovery Grant, NSERC, Canada.
[1]Departments of Biostatistics and Statistics & Operations Research, University of North Carolina at Chapel Hill, NC, U. S. A. e-mail: pksen@bios.unc.edu
[2]Department of Mathematics and Statistics, Carleton University, Ottawa, ON K1S5B6, Canada. e-mail: esaleh@math.carleton.ca
AMS 2000 subject classifications: Primary 62G05, 62G20; secondary 62G35
Keywords and phrases: asymptotic normality, contiguity of probability measures, discounted slope, divided differences, Kendall's tau, least squares estimator, median unbiasedness, uniform asymptotic linearity.

are rarely tenable. A Box-Cox type transformation on either variable may not only distort the form of the regression line but also complicate the error structures. Hence, it seems more reasonable to look into the usual ME model without the normality of errors or even the finiteness of the error variances; the latter aspect is important from robustness perspectives. The present study is mainly a characterization of the TSE in a ME setup where the normality of errors is dispensed with less stringent regularity assumptions. An important by-product of this characterization of TSE is the scope for indepth study of various finite sample to asymptotic properties of the TSE in a ME setup. Since the TSE is a member of a general class of (regression) R-estimators (which unlike the TSE may not have a closed expression), a formulation of the *contiguity of probability measures* in a ME setup is incorporated here to facilitate the study of asymptotic properties of such general nonlinear estimators. For the TSE, the contiguity based derivation of asymptotic properties is, however, not that essential, and under slightly different regularity conditions, a direct approach is presented along with.

In passing, we may remark that in the simple regression model, the TSE provides a distribution-free confidence interval for the slope β. This procedure (Sen [11]) rests on an independence clause whereby the permutation distribution of the Kendall tau statistic under the hypothesis of no regression agrees with its null distribution. In a ME setup, this simple equivariance result may not be generally true, and hence, alternative approaches are to be developed for the confidence interval problem.

2. Preliminary notion

Consider a simple regression (without ME) model with dependent variable Y_i and (nonstochastic) independent or explanatory variable t_i:

$$Y_i = \theta + \beta t_i + e_i, i = 1, \ldots, n; \tag{2.1}$$

where θ is the intercept parameter, β is the slope parameter, the e_i are independent and identically distributed (i.i.d.) error variables with mean zero and finite variance σ_e^2, and $t_1, \ldots, t_n$ are known regression constants, not all equal. In this setup, the LSE of slope parameter β can be expressed as

$$\hat{\beta}_n = \frac{\sum_{1\leq i<j\leq n}(Y_j - Y_i)(t_j - t_i)}{\sum_{1\leq i<j\leq n}(t_j - t_i)^2}. \tag{2.2}$$

Set $\mathcal{S} = \{1 \leq r < s \leq n : t_s \neq t_r\}$ and define *divided differences* and relative weights as

$$\begin{aligned} Z_{ij} &= (Y_j - Y_i)/(t_j - t_i), \\ &= \beta + (e_j - e_i)/(t_j - t_i) = \beta + Z_{ij}^o, \text{ say}, \\ w_{ij} &= (t_j - t_i)^2 \Big/ \sum_{1\leq r<s\leq n} (t_s - t_r)^2, \end{aligned} \tag{2.3}$$

for $(i, j) \in \mathcal{S}$. Then, we have

$$\begin{aligned} \hat{\beta}_n &= \sum_{\{(i,j)\in\mathcal{S}\}} w_{ij} Z_{ij} \\ &= \beta + \sum_{\{(i,j)\in\mathcal{S}\}} w_{ij} Z_{ij}^o. \end{aligned} \tag{2.4}$$

Thus, whenever e_i has a finite variance σ_e^2, even without the normality of the errors, $E\hat{\beta}_n = \beta$ and

$$\mathrm{Var}(\hat{\beta}_n) = \sigma_e^2 \Big/ \sum_{i=1}^{n} (t_i - \bar{t}_n)^2. \tag{2.5}$$

This representation reveals that the LSE is very sensitive to outliers and has low efficiency for heavy-tailed distributions, along with some other undesirable properties (Sen [11]). By contrast, the TSE of β is simply given by the median of the $Z_{ij}, (i,j) \in \mathcal{S}$ (Sen [11]). This estimator, basically being a median of some dependent, non-i.i.d. but symmetrically distributed divided differences, exhibits greater robustness for outliers, error contamination etc. Let us consider next a ME setup and appraise the extent to which the properties of LSE and TSE are compromised.

3. The ME model

Let us consider a motivating illustration. It is of interest to regress Y, the systolic blood pressure (SBP) on W, the body mass index (BMI). Even for the same person, the SBP is known to vary over time or other extraneous factors and is also subject to ME due to recording instrument. Likewise, the BMI is measured indirectly through other physiological measurements and is thereby subject to intrinsic as well as instrumental errors. As such consider an observable set of n independent stochastic vectors $(Y_i, W_i), i = 1, \ldots, n$ where

$$\begin{aligned} Y_i &= Y_i^o + \eta_i, W_i = X_i + U_i, \\ Y_i^o &= \mu_y + \beta X_i + e_i, X_i = \mu_x + V_i, i = 1, \ldots, n, \end{aligned} \tag{3.1}$$

and the error components U_i, V_i, e_i and η_i are mutually independent. Note that η_i and U_i are the measurement errors on the Y^o and X variables respectively, while e_i, V_i relate to intrinsic chance error for the unobservable Y_i^o, X_i. Here, η_i does not affect the regression but U_i has an affecting role in the regression. This model is known as *error in variables* (EIV) models, considered by Fuller [1] and others. The contemplated ME model is also known as the *simple structural linear relation model with model error*, and we refer to Hsiao [5] and Kukush and Zwanzig [8] where other pertinent references are cited.

When all the error components are assumed to be normally distributed (entailing finite variances $\sigma_e^2, \sigma_\eta^2, \sigma_u^2$ and σ_v^2), the regression of Y on W is linear with the slope parameter $\gamma = \kappa\beta$ where the discounting factor $\kappa,\ 0 \le \kappa \le 1$, is given by

$$\kappa = \sigma_v^2 / \{\sigma_v^2 + \sigma_u^2\}. \tag{3.2}$$

Further, in this normal error model, $Y - \gamma W$ and W are stochastically independent. This simple resolution may not workout when the errors are not all normally distributed: even if the the error variances are finite, $Y - \gamma W$ and W may be uncorrelated but not necessarily independent. Even the uncorrelation may not hold if the error variances are not finite.

Assuming the error variances to be finite, if we blindly use the LSE of Y on W it is given by

$$\hat{\gamma}_{nL} = \frac{\sum_{1 \le i < j \le n} (Y_i - Y_j)(W_i - W_j)}{\sum_{1 \le i < j \le n} (W_i - W_j)^2}. \tag{3.3}$$

Note that the LSE is a ratio of two U-statistics (Hoeffding [4]), and hence, under finite variances of the errors, it converges almost surely (a.s.) to γ as $n \to \infty$. Thus, normality of the errors is not crucial for the LSE to be (strongly) consistent for γ. However, without normality of errors, strictly unbiasedness or even median unbiasedness of the LSE may not hold. To gain further insight, we follow an *estimating equation* (EE) approach. Recall that (Y_i, W_i) are are i.i.d. stochastic vectors with $\mathrm{Cov}(Y_i, W_i) = \mathrm{Cov}(Y_i^o + \eta_i, X_i + U_i) = \mathrm{Cov}(Y_i^o, X_i) = \beta\sigma_v^2$. Thus, if we let for a given (real) b,

$$\begin{aligned} S_n(b) &= \sum_{1\le i<j\le n} (W_i - W_j)(Y_i - bW_i - Y_j + bW_j) \\ &= \sum_{1\le i<j\le n} (W_i - W_j)(Y_i - Y_j) - b \sum_{1\le i<j\le n} (W_i - W_j)^2. \end{aligned} \tag{3.4}$$

then $S_n(b)$ is a strictly monotone decreasing function of $b \in \mathcal{R}$. Further note that

$$E(W_i - W_j)^2 = 2(\sigma_u^2 + \sigma_v^2) = 2\sigma_w^2. \tag{3.5}$$

Hence, $E_\beta S_n(b) = 0$ only when $b = \beta\sigma_v^2/(\sigma_v^2 + \sigma_u^2) = \gamma$. Thus, the graph of $(b, S_n(b)), b \in \mathcal{R}$ crosses the abscissa at $b = \hat{\gamma}_{nL}$ which is the LSE.

For nonnormal errors, $S_n(\gamma)$ may not have a symmetric distribution around 0, and hence, the median-unbiasedness of the LSE may not hold. Also, since the LSE is the ratio of two U-statistics, it may not be unbiased for γ. However, by Theorem 7.5 of Hoeffding [4] under finite 4th order moments of all the errors, the asymptotic normality of the LSE follows readily. This result clearly depicts the high degree of nonrobustness of LSE to outliers, error contamination and its inefficiency for heavy-tailed distributions. Moreover, for nonnormal errors, the LSE may not provide an exact confidence interval for γ.

Motivated by this less than desired performance characteristics of the LSE in a ME setup, we intend to study the performance of the TSE. In passing, we may remark that η_i being independent of e_i, U_i, V_i can easily be absorbed in the e_i without affecting the relation with U_i, V_i, and hence, in the sequel, we omit η_i in the basic model (3.1) and work with the Y_i instead of the Y_i^o. Though this adjustment does not affect the estimation of the parameters, in the expression for their standard errors, η will add additional variability. The convoluted density of e_i and η_i takes care of that adjustment.

We follow the EE approach for the TSE too. As in Sen [11], we consider the following form of the aligned Kendall tau statistic, convenient to deal with in the contemplated ME model. For real $b \in \mathcal{R}$, we set

$$K_n(b) = \sum_{1\le i<j\le n} sign((Y_i - Y_j) - b(W_i - W_j))sign(W_i - W_j). \tag{3.6}$$

Since $sign(ab) = sign(a)sign(b)$, we rewrite $K_n(b)$ as

$$K_n(b) = \sum_{1\le i<j\le n} sign(Z_{ij} - b),\ b \in \mathcal{R}, \tag{3.7}$$

where $Z_{ij} = (Y_i - Y_j)/(W_i - W_j)$. As such, $K_n(b)$ is nonincreasing (and a step down function) in b.

The crux of the problem is therefore to study the nature of $E_\beta Z_{ij} - b$ in a ME model and develop an estimating equation accordingly.

4. Rationality of TSE in ME model

Let us define $U_{ij} = U_i - U_j,\ V_{ij} = V_i - V_j,\ e_{ij} = e_i - e_j$, so that we have

$$\begin{aligned} Y_i - Y_j - b(W_i - W_j) &= e_{ij} + \beta V_{ij} - b(U_{ij} + V_{ij}) \\ &= e_{ij} + (\beta - b)V_{ij} - bU_{ij}, \end{aligned} \tag{4.1}$$

for all $1 \leq i < j \leq n$. Recall that e_{ij}, U_{ij}, V_{ij} are all independent and each one has a symmetric distribution around 0. However, this symmetry is not enough to guarantee the desired pivotal result. We denote the density function of the e_{ij}, U_{ij} and V_{ij} by $f_e(\cdot), f_u(\cdot)$ and $f_v(\cdot)$ respectively. While we allow $f_e(\cdot)$ to be completely arbitrary but symmetric about 0, for the other two densities, in view of their symmetric form around 0, we make the following **Assumption A**, linking them to a common member of the location-scale family of densities; the conventional normal case is a particular one in this general family:

$$\begin{aligned} f_u(x) &= \lambda_u^{-1} f_o(x/\lambda_u), \\ f_v(x) &= \lambda_v^{-1} f_o(x/\lambda_v), \end{aligned} \tag{4.2}$$

where $f_o(\cdot)$ is a symmetric density free from nuisance parameter(s) and λ_u, λ_v are unknown scale parameters.

If we assume that the density $f_o(\cdot)$ admits of a finite variance say σ_o^2 then

$$var(U_{ij}) = \lambda_u^2 \sigma_o^2;\ var(V_{ij}) = \lambda_v^2 \sigma_o^2. \tag{4.3}$$

The last two equations also imply that $U_{ij}^* = U_{ij}/\lambda_u$ and $V_{ij}^* = V_{ij}/\lambda_v$ both have the common density $f_o(\cdot)$ and hence are identically distributed; this feature remains in tact even if σ_o does not exist. Further, whenever σ_o^2 is finite, we note that

$$\kappa = \frac{\sigma_v^2}{\sigma_v^2 + \sigma_u^2} = \frac{\lambda_v^2}{\lambda_v^2 + \lambda_u^2}. \tag{4.4}$$

Henceforth, we shall express γ in terms of λ_u and λ_v.

In the above setup, if we let $b = \gamma = \kappa\beta$ then

$$\begin{aligned} e_{ij} &+ (\beta - \gamma)V_{ij} - \gamma U_{ij} \\ &= e_{ij} + \beta[(1 - \kappa)V_{ij} - \kappa U_{ij}], \end{aligned} \tag{4.5}$$

where e_{ij} is independent of both U_{ij}, V_{ij}. Further, $V_{ij} = (\lambda_v)V_{ij}^*$ has the same (symmetric) density as $(\lambda_v)U_{ij}^* = [U_{ij}(\lambda_v/\lambda_u)]$. Moreover, note that $(1 - \kappa)/\kappa = \lambda_u^2/\lambda_v^2$ so that

$$\begin{aligned} V_{ij}/\lambda_v &- [\kappa/(1 - \kappa)](U_{ij}^*)[\lambda_u/\lambda_v] \\ &= V_{ij}^* - [\kappa/(1 - \kappa)]^{1/2} U_{ij}^*. \end{aligned} \tag{4.6}$$

Further, noting that U_{ij}^* and V_{ij}^* are i.i.d. both having a common symmetric density $f_o(\cdot)$, we conclude that the joint density of (U_{ij}^*, V_{ij}^*) is totally symmetric around the origin $\mathbf{0}$. As such, if we let

$$\begin{aligned} L_{ij} &= V_{ij}^* - \sqrt{\kappa/(1 - \kappa)}\, U_{ij}^*, \\ Q_{ij} &= U_{ij}^* + \sqrt{\kappa/(1 - \kappa)}\, V_{ij}^*, \end{aligned} \tag{4.7}$$

we may express $(L_{ij}, Q_{ij}) = (U^*_{ij}, V^*_{ij})\mathbf{P}$, where $\sqrt{(1-\kappa)}\mathbf{P}$ is an orthogonal matrix. Therefore, invoking the total symmetry of the joint density of (U^*_{ij}, V^*_{ij}), we conclude that (L_{ij}, Q_{ij}) has a totally symmetric joint density around the origin. If the independent U^*_{ij}, V^*_{ij} were normally distributed, L_{ij}, Q_{ij} would have been independent too. However, sans the normality of the U^*_{ij}, V^*_{ij}, the L_{ij}, Q_{ij} would be uncorrelated but not necessarily independent. Hence, this characterization of total symmetry of the joint distribution of L_{ij}, V_{ij} is the best we could get and that serves our purpose too. Next, we note that

$$\begin{aligned} U_{ij} + V_{ij} &= \sqrt{(1-\kappa)/\kappa}\lambda_v U^*_{ij} + \lambda_v V^*_{ij} \\ &= \lambda_v \sqrt{(1-\kappa)/\kappa} Q_{ij} \\ &= \lambda_u Q_{ij}. \end{aligned} \tag{4.8}$$

and at $b = \kappa\beta$,

$$\begin{aligned} e_{ij} &+ \beta[(1-\kappa)\lambda_v V^*_{ij} - \kappa\lambda_u U^*_{ij}] \\ &= e_{ij} + \beta\lambda_v(1-\kappa)L_{ij} \end{aligned} \tag{4.9}$$

has a symmetric distribution around 0. This is also a linear combination of e_{ij} and L_{ij} (which are independent), and L_{ij} is orthogonal to Q_{ij}. Thus, we conclude that for any combination $(e_{ij}, L_{ij}, Q_{ij}) = (e, l, q)$, we can define an *orbit* O of 16 mass points: $(e, l, q), (-e, l, q), (e, -l, q), (e, l, -q), (-e, -l, q), (-e, l, -q), (e, -l, -q), (-e, -l, -q), (e, q, l), (-e, q, l), (e, -q, l), (-e, -q, l), (e, q, -l), (-e, -q, l), (e, -q, -l), (-e, -q, -l)$ such that the conditional distribution of $(e_{ij}.L_{ij}, Q_{ij})$ on this orbit is discrete uniform with a (conditional) probability mass $1/16$ attached to each of these 16 points. Of these 16 points, 8 lead to $+1$ and remaining 8 to -1 for the kernel. Therefore, first taking conditional expectation over an orbit and then integrating over all orbits, it can be concluded that under Assumption A,

$$E_\beta\{K_n(\gamma)\} = 0. \tag{4.10}$$

Along with this result, the monotonicity of $E_\beta[K_n(b)]$ in b provide the rationality of the estimating equation $K_n(b) = 0$ which yields the TSE of γ in the ME model. As such, the TSE is denoted by

$$\hat{\gamma}_{nT} = \text{median}\{Z_{ij} : (i,j) \in \mathcal{S}\}. \tag{4.11}$$

We may also set $\hat{\gamma}_{nT} = \gamma + \text{med}\{e_{ij} + \beta\lambda_v(1-\kappa)L_{ij} : (i,j) \in \mathcal{S}\}$, where in view of the stochastic nature of the W_i and their continuous distributions, $\mathcal{S}$ can be replaced by the set of all $\binom{n}{2}$ pairs $(1 \le i < j \le n)$. Further, note that $K_n(\gamma+\epsilon)/\binom{n}{2} \to \delta(\epsilon)$ a.s., as $n \to \infty$, where $\delta(\epsilon)$ is negative or positive according as ϵ is positive or negative. This result follows from the a.s. convergence of U-statistics. Hence, we arrive at the main result of this section.

Theorem 4.1. Under Assumption A, the estimating equation $K_n(b) = 0$ leads to the TSE $\hat{\gamma}_{nT}$ which is a strongly consistent estimator of γ.

5. Median-unbiasedness of TSE

Note that $K_n(b)$ is invariant under any any change of μ_y, μ_x, and hence, without any loss of generality, we set $\mu_y = \mu_x = 0$. As such, for $K_n(\gamma)$, we work with the variables $(e_i + \beta[(1-\kappa)V_i - \kappa U_i], U_i + V_i) = (L_i, Q_i)$, say $i = 1, \ldots, n$. We denote by

$$K_n(\gamma) = K_n((L_1, Q_1), \ldots, (L_n, Q_n)). \tag{5.1}$$

Then, by arguments (on total symmetry) similar to the preceding section, we claim that under Assumption A,

$$K_n((L_1, -Q_1), \ldots, (L_n, -Q_n)) = -K_n((L_1, Q_i), \ldots, (L_n, Q_n)), \tag{5.2}$$

so that the distribution of $K_n(\gamma)$ is symmetric about 0. This, in turn implies that

$$\begin{aligned} P_\beta\{\hat{\gamma}_{nT} \le \gamma\} &= P_\beta\{K_n(\gamma) \ge 0\} \\ &= P_\beta\{K_n(\gamma) \le 0\} = P_\beta\{\hat{\gamma}_{nT} \ge 0\}. \end{aligned} \tag{5.3}$$

so that the TSE is median-unbiased for γ. In the above derivation of median-unbiasedness of TSE, we have tacitly bypassed the role of finite variances of e_i, U_i, V_i, and hence, the results pertain to a general class of densities, including the Cauchy, where the variances may not necessarily exist. We may also remark that the (Y_i, W_i), $i \ge 1$, are i.i.d. stochastic vectors, and hence, for every $(i,j) \in \mathcal{S}$, Z_{ij} has a symmetric distribution; we denote this common distribution by $G(z), z \in \mathcal{R}$. Using then the moment properties of sample quantiles, as extended to U-processes, it can be shown that if $G(\cdot)$ admits of a finite absolute moment of order δ for some $\delta > 0$, not necessarily an integer, then for every $n \ge 4k/\delta$, the TSE has a finite (absolute) moment of order k. Hence, for $n \ge 4/\delta$, TSE is unbiased for the discounted slope parameter γ. For i.i.d.r.v, this moment result of sample quantile is due to Sen [10], and the rest of the proof follows by noting that the tail probability of the TSE is dominated by the tail probability of median$\{Z_{12}, \ldots, Z_{2m-1,2m}\}$ where m is the largest integer contained in $(n+1)/2$.

6. General asymptotics of TSE

Here, in the ME setup, we discuss the asymptotic results without incorporating contiguity of probability measures. Note that the kernel in the definition of $K_n(b)$ is bounded so that moments of all finite order exist. Because of the non-increasing (step-down) property of $K_n(b)$, $b \in \mathcal{R}$, and the boundedness of the kernel in the Kendall tau statistic, the asymptotic normality and some other properties of TSE are studied by relatively simpler and direct analysis, along the lines in Section 4.

First, note that $E_\beta K_n(b)$ is a continuous and monotone decreasing function of $b \in \mathcal{R}$. Further, if we set $b = b_n = \gamma + n^{-1/2}\xi$, for some fixed ξ, then for any pair (i,j),

$$\begin{aligned} e_{ij} \; &+ \; (\beta - b_n)V_{ij} - b_n U_{ij} \\ &= e_{ij} + \beta\lambda_v(1-\kappa)L_{ij} - \frac{\xi\lambda_v\sqrt{1-\kappa}}{\sqrt{n\kappa}}Q_{ij} \\ &= e_{ij} + \beta\lambda_v(1-\kappa)L_{ij} - n^{-1/2}\xi\lambda_u Q_{ij}, \end{aligned} \tag{6.1}$$

where the e_{ij}, L_{ij}, Q_{ij} are all defined in Sections 1−4. In the following, for simplicity, we let $\xi > 0$ (and a similar treatment holds for $\xi < 0$). As such, if we consider a specific pair (i,j) in the summand of $K_n(\gamma + n^{-1/2}\xi)$, its expectation comes out as

$$\begin{aligned} -4P\{Q_{ij} > \;\; 0; \;\; -\beta(1-\kappa)\lambda_v L_{ij} \; &\le e_{ij} \\ &\le -\beta(1-\kappa)\lambda_v L_{ij} + n^{-1/2}\xi\lambda_u Q_{ij}\}. \end{aligned} \tag{6.2}$$

We denote the joint distribution function of (L_{ij}, Q_{ij}) by $H^*(l,q)$, $(l,q) \in \mathcal{R}^2$. Also, as in Section 4, we denote the density of e_{ij} by $f_e(\cdot)$ Further, assume that

Assumption B: The following functional exists:

$$A^* = \int_{\mathcal{R}} \int_0^\infty q f_e(\beta(1-\kappa)\lambda_v l)\, \mathrm{d}H^*(l,q). \tag{6.3}$$

It is easy to show that

$$\lambda_u A^* = (1/2)E\{f_e(\beta((1-\kappa)V_{ij} - \kappa U_{ij})|U_{ij} + V_{ij}|\}. \tag{6.4}$$

In the case of no measurement error, $\lambda_u = 0$ and $U_i = 0$ (a.e.), and hence, $\kappa = 1$, so that the last expression reduces to

$$(1/2)f_e(0)E|V_{ij}|. \tag{6.5}$$

Even this expression is different from the case where the x_i are nonstochastic, as treated in Section 2. Further, note that

$$f_e(0) = \int_{\mathcal{R}} f_e^{*2}(e)\, \mathrm{d}e, \tag{6.6}$$

where $f_e^*(\cdot)$ is the pdf of e_i. In passing, we may remark that a sufficient condition for A^* to be finite is that $H^*(\cdot)$ admits of a finite first order moment and $f_e(\cdot)$ is bounded a.e. A less restrictive condition would be to assume the integrability of $Q_{ij} f_e(\beta(1-\kappa)\lambda_v L_{ij})$. Then, by standard manipulations, along the lines of Section 4, it follows that (6.2) is asymptotically

$$-2n^{-1/2}A^*\xi\lambda_u + o(n^{-1/2}). \tag{6.7}$$

Next, we note that for any fixed ξ,

$$P\{\sqrt{n}(\hat{\gamma}_{nT} - \gamma) \le \xi\} \le P\{\sqrt{n}K_n(\gamma + n^{-1/2}\xi) \le 0\}, \tag{6.8}$$

and a lower bound to the left hand side of (6.5) is the right hand side with $\le$ being replaced by strict inequality (< 0). As such, for large n, we can work with either the upper or lower bound in (6.5). Since, for any b, $K_n(b)$ is a U-statistic based on a bounded kernel of degree 2, its asymptotic normality holds with appropriate mean and variance functions. Since, here $b = b_n = \gamma + \xi n^{-1/2}$, the asymptotic variance can be replaced by the corresponding expression at $b = \gamma$ but the mean has to be adjusted according to (6.2). As such, (6.8) is asymptotically equivalent to

$$P\{\sqrt{n}[K_n(\gamma + \xi/\sqrt{n}) - EK_n(\gamma + \xi/\sqrt{n})] \le 2\xi\lambda_u A^*\}. \tag{6.9}$$

Further, note that as $n \to \infty$,

$$n\mathrm{Var}(K_n(\gamma)) \to 4\nu^2, \tag{6.10}$$

where ν^2 is the variance of the first order kernel corresponding to the kernel of $K_n(\gamma)$ (Hoeffding [4]).

We need to address ν^2 a bit more elaborately than in the conventional regression model, treated in Section 2. Note that (Y_i, W_i) are i.i.d. random vectors, and hence, $Y_i^* = Y_i - \gamma W_i, i = 1, \ldots, n$ are i.i.d.r.v.. Therefore $Y_i^* - Y_j^*$ has a symmetric distribution around 0. On the other hand, in the ME setup, as has been discussed earlier, $Y_i^* - Y_j^*$ and $W_i - W_j$ are not generally independent (but are uncorrelated); they are independent in the case where the errors U_i, V_i are normally distributed (irrespective of the distribution of e_{ij}). Keeping this in mind, we denote the joint distribution function of (Y_i^*, W_i) by $H(y^*, w)$, for $(y^*, w) \in \mathcal{R}^2$. Then we note that

$$\begin{aligned} & E[sign((Y_i^* - Y_j^*)(W_i - W_j))|Y_i^*, W_i] \\ = \ & 4H(Y_i^*, W_i) - 2H_i(Y_i^*) - 2H_2(W_i) + 1, \end{aligned} \tag{6.11}$$

where $H_1(\cdot)$ and $H_2(\cdot)$ refer to the marginal distribution functions. Further, note that by arguments presented in Section 4,

$$E[4H(Y_i^*, W_i) - 2H_1(Y_i^*) - 2H_2(W_i) + 1] = 0. \tag{6.12}$$

As a result, we obtain that

$$\begin{aligned} \nu^2 &= E\{[4H(Y_i^*, W_i) - 2H_1(Y_i^*) - 2H_2(W_i) + 1]^2\} \\ &= \int\int_{\mathcal{R}^2} [4H(y, w) - 2H_1(y) - 2H_2(w) + 1]^2 \, \mathrm{d}H(y, w). \end{aligned} \tag{6.13}$$

Note that when Y_i^*, W_i are independent, $H(y, w) = H_1(y)H_2(w)$, and hence, the above expression reduces to 1/9, so that $4\nu^2 = 4/9$, the leading term in the variance of $\sqrt{n}K_n(0)$ under the null hypothesis of independence of Y_i^*, W_i.

Having checked the expression (6.13) for ν^2 in a general ME setup, and appealing to the celebrated theorem of Hoeffding [4] on the asymptotic normality of a U-statistic when the parameter is stationary of order 0, we complete the proof of asymptotic normality of the TSE in ME model by using (6.8) and (6.9). Hence, we have the following.

Theorem 6.1 . Under Assumptions (A,B), for every fixed $\xi \in \mathcal{R}$, as $n \to \infty$,

$$P\{\sqrt{n}(\hat{\gamma}_{nT} - \gamma) \leq \xi\} \to \Phi(\xi/\zeta), \tag{6.14}$$

where $\Phi(x), x \in \mathcal{R}$ is the standard normal distribution function and

$$\begin{aligned} \zeta^2 &= \nu^2\kappa/\{\lambda_v A^*(1-\kappa)\}^2 \\ &= \nu^2/\{A^{*2}\lambda_u^2\}. \end{aligned} \tag{6.15}$$

The last result yields, as a special case, the asymptotic normality of the TSE in the nonstochastic regressor case as treated in Sen [11] and elsewhere, albeit the expression for ν^2 could be different as the regressors are not necessarily distinct.

7. Contiguity in ME models

We conclude this study with a general observation on the contiguity of probability measures in the ME model in Section 3; this result pertains to general linear rank statistics as well as other likelihood based ones. In the hypothesis testing context, a similar result for (partially informed) stochastic regressors was established by Ghosh and Sen [2]). More recently, Jurečková, Picek and Saleh [6] studied the testing problem in a ME setup using regression rank scores. Also, Saleh, Picek and Kalina [9] have studied nonparametric estimation in ME models, putting major emphasis on numerical studies. Under the ME setup, the verification of contiguity is simpler and neater too. Further, in view of the monotonicity of $K_n(b)$ in $b \in \mathcal{R}$, the *uniform asymptotic linearity results* presented in detail in Jurečková and Sen [7] may not be needed in this specific case.

We use the same notation as in Section 3, and note that the observable r.v.s $(Y_i, W_i),\ i = 1, \ldots, n$ are identically distributed. We denote the (bivariate) density function of (Y_i, W_i) by $f_{Y,W}(y, w), (y, w) \in \mathcal{R}^2$. Also, let $f_X(x), x \in \mathcal{R}$ be the marginal density of X_i (unobservable). Then, we can write

$$f_{Y,W}(y, w) = \int_{\mathcal{R}} f(y, w|x) f_X(x)\, \mathrm{d}x. \tag{7.1}$$

Next, we write $f(y, w|x) = f(y|w, x)f(w|x)$. At this stage, WLOG, we take $\mu_y = 0 = \mu_x$, and note that given W, X, the conditional density of Y depends only on X. This along with (3.1) lead to

$$f_{Y,W}(y, w) = \int_{\mathcal{R}} f_e(e - \beta v) f_U(w - v) f_V(v)\, \mathrm{d}v, \tag{7.2}$$

where $y = \beta v + e, x = v, w = u + v$. Therefore, we have

$$\begin{aligned} &(\partial/\partial\beta) f_{e,w}(e, u + v; \beta) \\ &= \int_{\mathcal{R}} [(\partial/\partial\beta) \log f_e(e - \beta v)] f_e(e - \beta v) f_U(w - v) f_V(v)\, \mathrm{d}v, \end{aligned} \tag{7.3}$$

provided the usual regularity conditions which permit the interchange of integration (over v) and differentiation (with respect to β) hold. Further, note that the partial derivative (wrt β) inside the above integral is equal to $-v(\partial/\partial e) \log f_e(e-\beta v)$, and we write this as $v\psi(e - \beta v)$, where

$$\psi(e - \beta v) = -(\partial/\partial e) f_e(e - \beta v)/f_e(e - \beta v) \tag{7.4}$$

is the usual Fisher score function associated with the density $f_e(\cdot)$. Also, note that

$$f(v|e, w) = \frac{f_e(e - \beta v) f_U(w - v) f_V(v)}{\int_{\mathcal{R}} f_e(e - \beta v) f_U(w - v) f_V(v)\, \mathrm{d}v}. \tag{7.5}$$

As a result, $(\partial/\partial\beta) \log f_{e,w}(e, w; \beta)$ can be written as

$$\psi^*(e, w) = \int_{\mathcal{R}} v\psi(e - \beta v) f(v|e, w)\, \mathrm{d}v. \tag{7.6}$$

Thus, it is easy to show that the expected value of the left hand side of (7.6) is equal to 0 (as it should be). Therefore, under the usual (Cramér) regularity conditions on the pdf $f_e(\cdot), f_U(\cdot)$ and $f_V(\cdot)$ along with the following:

Assumption C: $\psi^*(e, w)$ is square integrable.
It is easy to verify contiguity by invoking Le Cam's First and Second lemma (viz., Hájek et al. [3], ch. 7). Moreover, using the Jensen inequality along with the Cauchy–Schwarz inequality, it follows that

$$\begin{aligned} E[(\partial/\partial\beta) \log f_{e,w}(e, w; \beta)]^2 &= E\{[E(v\psi(e - \beta v)|e, w)]^2\} \\ &\leq E(V^2) E(\psi^2(e - \beta V)) \end{aligned} \tag{7.7}$$

so that the finite fisher information of the score function associated with the pdf $f_e(\cdot)$ along with the finite second moment of W will provide a set of sufficient conditions. The technical details are therefore omitted.

In passing we may remark that for the TSE based on the Kendall tau statistic having a bounded kernel, the contiguity based proof of asymptotic normality is not needed, and the needed regularity Assumptions A, B are relatively less restrictive than C. However, the last expression conveys an easily verifiable condition, albeit under the finite variance of the regressor; for $f_e(\cdot)$ the finite Fisher information suffices. In Assumption B, the finiteness of the variance of V is not needed. For general linear rank statistics based R-estimator in a general ME model, the underlying score generating function may not be bounded, and we may not have a closed expression for the estimator. In such a case, the contiguity based proof of asymptotic normality should be a more plausible approach. We intend to pursue this in a subsequent communication.

Acknowledgements. We are grateful to both the reviewers for their most useful comments on the original draft. Their painstaking reading has eliminated a large number of typos and improved the presentation as well.

References

[1] Fuller, W. (1987). *Measurement Error Models.* John Wiley, New York.

[2] Ghosh, M. and Sen, P. K. (1971). On a class of rank order tests for regression with partially informed stochastic predictors. *Annals of Mathematical Statistics* **42** 650–661.

[3] Hájek, J., Šidák, Z., and Sen, P. K. (1999). *Theory of Rank Tests.* 2nd Ed., Academic Press, London.

[4] Hoeffding, W. (1948). On a class of statistics with asymptotically normal distribution. *Annals of Mathematical Statistics* **19** 293–325.

[5] Hsiao, Cheng (1989). Consistent estimation for some nonlinear errors-in-variables models. *Journal of Econometrics* **41** 159–185.

[6] Jurečková, J. Picek, J., and Saleh, A. K. M. E. (2009). Rank tests and regression rank score tests in measurement error models. *Computational Statistics and Data Analysis*, in press.

[7] Jurečková, J. and Sen, P. K. (1996). *Robust Statistical Procedures: Asymptotics and Interrelations.* John Wiley, New York.

[8] Kukush, A. and Zwanzig, S. (2002). On consistent nonlinear estimation in functional errors in variables models. In *Total Least Squares and Errors-in-Variable Modelling* (eds. Huffel, S. and Lemmerling, P.), Kluwer, Boston.

[9] Saleh, A. K. M. E., Picek, J., and Kalina, J. (2009). Nonparametric estimation of regression parameters in measurement error models. *Metron* **LXVII** 177–200.

[10] Sen, P. K. (1959). On the moments of sample quantiles. *Calcutta Statistical Association Bulletin* **9** 1–19.

[11] Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association* **63** 1379–1389.