The Annals of Statistics
1996, Vol. 24, No. 4, 1536–1546

ASYMPTOTICALLY HONEST CONFIDENCE SETS FOR STRUCTURAL ERRORS-IN-VARIABLES MODELS[1]

BY LONGCHEEN HUWANG

Cornell University and Tsing Hua University

The problem of constructing confidence sets for the structural errors-in-variables model is considered under the assumption that the variance of the error associated with the covariate is known. Previously proposed confidence sets for this model suffer from the problem that they all have zero confidence levels for any sample size, where the confidence level of a confidence set is defined to be the infimum of coverage probability over the parameter space. In this paper we construct some asymptotically honest confidence sets; that is, the limiting values of their confidence levels are at least as large as the nominal probabilities when the sample size goes to ∞. A desirable property of the proposed confidence set for the slope is also established.

1. Introduction. Suppose that there are unobservable "true" random variables (u_i, η_i) that satisfy a linear relation,

$$\eta_i = \alpha + \beta u_i. \tag{1.1}$$

We can only observe (X_i, Y_i) which are the true random variables plus additive measurement errors $(\delta_i, \varepsilon_i)$; that is,

$$Y_i = \eta_i + \varepsilon_i, \qquad X_i = u_i + \delta_i, \qquad i = 1, \ldots, n, \tag{1.2}$$

where

$$(u_i, \delta_i, \varepsilon_i)' \sim_{\text{i.i.d.}} N[(m_u, 0, 0)', \operatorname{diag}(\sigma_u^2, \sigma_\delta^2, \sigma_\varepsilon^2)],$$

$-\infty < \alpha, \beta, m_u < \infty$, σ_u^2, σ_δ^2 and $\sigma_\varepsilon^2 > 0$. Model (1.1)–(1.2) is called a structural errors-in-variables model since the covariates u_i are independent, identically distributed random variables. This model is different from a functional errors-in-variables model, where u_i are assumed to be unknown constants, not random variables. Without extra assumptions about the parameters, however, the model is unidentifiable. To avoid this difficulty, typically it is assumed that one of the error variances or their ratio is known. Perhaps the most realistic assumption is that one of the error variances is estimable by an independent estimator, for example, when independent repeated observations made on u_i are available. Surveys of results for structural and functional errors-

Received August 1992; revised October 1995.

[1]Supported by NSF Grant DMS-88-09016.

AMS 1991 *subject classifications*. Primary 62F25; secondary 62J99, 62E99.

Key words and phrases. Errors-in-variables, confidence level, asymptotically honest confidence set, converge uniformly in all parameters.

in-variables models can be found in Moran (1971), Kendall and Stuart (1979), Anderson (1984) and Fuller (1987).

Let $\mathbf{X} = (X_1, \dots, X_n)$, $\mathbf{Y} = (Y_1, \dots, Y_n)$ and $C(\mathbf{X}, \mathbf{Y})$ be a confidence set for β [or (α, β)]. From now on, the (finite-sample) confidence level of the confidence set $C(\mathbf{X}, \mathbf{Y})$ is defined to be

$$\inf_{\boldsymbol{\theta} \in \Omega} P_{\boldsymbol{\theta}}[\beta \in C(\mathbf{X}, \mathbf{Y})],$$

where $\boldsymbol{\theta}$ and Ω denote the vector parameter and the corresponding parameter space, respectively. For the case $\sigma_\varepsilon^2/\sigma_\delta^2$ known, a confidence set for β with the confidence level exactly equal to the nominal probability has been previously derived by Creasy (1956). Similar results have, however, not been obtained for other types of identifiability assumptions, although confidence sets with asymptotic coverage probabilities equal to the nominal probabilities have been derived (see Section 2 for details). Here the asymptotic coverage probability of the confidence set $C(\mathbf{X}, \mathbf{Y})$ for β is defined as

$$\lim_{n \to \infty} P_{\boldsymbol{\theta}}[\beta \in C(\mathbf{X}, \mathbf{Y})].$$

Interestingly, the confidence level of a confidence set could be much smaller than the asymptotic coverage probability of the same confidence set. This seems to cause serious concern since the coverage probability of the confidence set on a certain subset of the parameter space could be dangerously low.

The results of Gleser and Hwang (1987) form a key part of the motivation for the present study. For models such as (1.1)–(1.2) their general theorem shows that any almost surely finite-diameter confidence set has a zero confidence level, where the diameter of a confidence set is defined to be the supremum distance between any two points in this set. As it turns out, the existing (except Creasy's) confidence sets have almost surely finite diameters and hence zero confidence levels no matter how large the sample size is. In this paper asymptotically honest confidence sets for β and (α, β) are constructed in the case when the error variance σ_δ^2 is known. Here a $100(1-\gamma)\%$ confidence set $C(\mathbf{X}, \mathbf{Y})$ for β is defined to be asymptotically honest [Li (1989)] if

$$\lim_{n \to \infty} \inf_{\boldsymbol{\theta} \in \Omega} P_{\boldsymbol{\theta}}[\beta \in C(\mathbf{X}, \mathbf{Y})] \geq 1 - \gamma, \tag{1.3}$$

where $\boldsymbol{\theta} = (\alpha, \beta, m_u, \sigma_u^2, \sigma_\varepsilon^2)$ and $\Omega = \{\boldsymbol{\theta};\ -\infty < \alpha, \beta, m_u < \infty,\ \sigma_u^2 > 0,\ \sigma_\varepsilon^2 > 0\}$. In fact, results stronger than asymptotic honesty are established in Sections 2 and 3. It is shown that the coverage probabilities of the proposed confidence sets converge to the nominal probabilities uniformly over the parameter space as the sample size goes to ∞. Furthermore, due to Gleser and Hwang's theorem, an asymptotically honest confidence set will inevitably have a positive probability of having an infinite diameter. However, among all asymptotically honest confidence sets for β, the proposed one minimizes the limit of the supremum as well as the limit of the infimum of this probability over the parameter space (see Theorem 2.5).

Although we do not address the more practical situation where the error variance σ_δ^2 is unknown but estimable [Madansky (1959)], the present results do have an implication for such a case. In fact, by substituting an independent estimator for σ_δ^2 and by taking into account the extra variability due to this substitution, the resultant confidence sets are also asymptotically honest. These results were established in Huwang (1991) and will be reported elsewhere.

The rest of the article is organized as follows. Section 2 deals with the confidence set for β, whereas Section 3 deals with the confidence set for (α, β). Difficulty in constructing the confidence set for α is discussed in Section 4.

2. Confidence set for the slope. Let $\bar{X}$, $\bar{Y}$, S_X^2, S_Y^2 and S_{XY}, respectively, denote $\sum X_i/n$, $\sum Y_i/n$, $\sum(X_i-\bar{X})^2/n$, $\sum(Y_i-\bar{Y})^2/n$ and $\sum(X_i - \bar{X})(Y_i-\bar{Y})/n$ throughout the paper. If the consistent estimators of α and β are defined by

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X}, \qquad \hat{\beta} = \frac{S_{XY}}{S_X^2 - \sigma_\delta^2},$$

it is well known [Fuller (1987), Theorem 1.2.1] that $\sqrt{n}(\hat{\alpha}-\alpha, \hat{\beta}-\beta)$ converges in distribution to a normal vector random variable with zero mean and covariance matrix

$$\Gamma = \begin{bmatrix} m_u^2\sigma_u^{-4}(\sigma_X^2\sigma_v^2 + \sigma_{Xv}^2) + \sigma_v^2 & -m_u\sigma_u^{-4}(\sigma_X^2\sigma_v^2 + \sigma_{Xv}^2) \\ -m_u\sigma_u^{-4}(\sigma_X^2\sigma_v^2 + \sigma_{Xv}^2) & \sigma_u^{-4}(\sigma_X^2\sigma_v^2 + \sigma_{Xv}^2) \end{bmatrix}, \tag{2.1}$$

where $v_i = Y_i - \alpha - \beta X_i = \varepsilon_i - \beta\delta_i$, σ_X^2 and σ_v^2 are the variances of X_i and v_i and σ_{Xv} is the covariance of X_i and v_i. Denote

$$V(\hat{\beta}) = \frac{1}{\sigma_u^4}(\sigma_X^2\sigma_v^2 + \sigma_{Xv}^2) = \frac{1}{\sigma_u^4}[\beta^2(\sigma_\delta^4 + \sigma_\delta^2\sigma_X^2) + \sigma_\varepsilon^2\sigma_X^2].$$

Then $\hat{\Gamma}$ converges in probability to Γ, where

$$\hat{\Gamma} = \begin{bmatrix} \bar{X}^2\hat{V}(\hat{\beta}) + S_v^2 & -\bar{X}\hat{V}(\hat{\beta}) \\ -\bar{X}\hat{V}(\hat{\beta}) & \hat{V}(\hat{\beta}) \end{bmatrix},$$

$$\hat{V}(\hat{\beta}) = \frac{1}{\hat{\sigma}_u^4}(S_X^2 S_v^2 + \hat{\beta}^2\sigma_\delta^4), \qquad S_v^2 = \frac{1}{n-2}\sum[Y_i - \bar{Y} - (X_i - \bar{X})\hat{\beta}]^2, \tag{2.2}$$

$$\hat{\sigma}_u^2 = S_X^2 - \sigma_\delta^2, \qquad \hat{\sigma}_{Xv} = -\hat{\beta}\sigma_\delta^2.$$

From this, it follows that, as $n \to \infty$,

$$T_n = \frac{\sqrt{n}(\hat{\beta}-\beta)}{\sqrt{\hat{V}(\hat{\beta})}} \to N(0, 1)$$

in distribution. In practice, it seems reasonable to approximate the distribution of T_n with the distribution of t with $n-2$ degrees of freedom. Consequently, a $100(1-\gamma)\%$ confidence set for β is defined by

$$|T_n| \le t_{n-2(\gamma/2)}, \tag{2.3}$$

where $t_{n-2(\gamma/2)}$ is the upper $100(\gamma/2)$ percentile of the t distribution with $n-2$ degrees of freedom. Although the confidence set (2.3) has a $1-\gamma$ asymptotic coverage probability, it has a finite diameter (almost surely) and hence a zero confidence level by virtue of Gleser and Hwang's theorem. Actually, by consecutively applying the dominated convergence theorem twice, it is easy to show that, for each n,

$$\lim_{\beta\to\infty} \lim_{\sigma_u^2\to 0} P_{\boldsymbol{\theta}}[|T_n| \le t_{n-2(\gamma/2)}] = 0.$$

Hence the finite-sample coverage probability of (2.3) could be extremely low on certain subsets of the parameter space Ω.

To derive an alternative confidence set for β, we consider the pivotal

$$T_n^* = \frac{\sqrt{n}(\hat{\beta}-\beta)}{\sqrt{V^*(\hat{\beta})}}, \tag{2.4}$$

where $V^*(\hat{\beta})$ is obtained from $V(\hat{\beta})$ wherein the parameters σ_u^2, σ_X^2 and σ_ε^2 are replaced by

$$\hat{\sigma}_u^2 = S_X^2 - \sigma_\delta^2, \qquad \hat{\sigma}_X^2 = S_X^2, \qquad \hat{\sigma}_\varepsilon^2 = \frac{n}{n-2}\left(S_Y^2 - \frac{S_{XY}^2}{S_X^2-\sigma_\delta^2}\right), \tag{2.5}$$

respectively. Precisely,

$$V^*(\hat{\beta}) = \frac{1}{\hat{\sigma}_u^4}[\beta^2(\sigma_\delta^4 + \sigma_\delta^2\hat{\sigma}_X^2) + \hat{\sigma}_\varepsilon^2\hat{\sigma}_X^2]. \tag{2.6}$$

Based on T_n^* in (2.4), a $100(1-\gamma)\%$ confidence set for β is defined by

$$|T_n^*| \le t_{n-2(\gamma/2)}. \tag{2.7}$$

Note that the major difference between $V^*(\hat{\beta})$ and $\hat{V}(\hat{\beta})$ (and hence T_n^* and T_n) is that, unlike $\hat{V}(\hat{\beta})$, the parameter β is not estimated in $V^*(\hat{\beta})$. Hence $V^*(\hat{\beta})$ here is not a real estimator of $V(\hat{\beta})$ (because it depends on β) but a "pseudo-estimator," with which we can form a pivotal to construct a confidence set. The approach of not estimating β in $V^*(\hat{\beta})$ has been previously proposed by Hwang (1988) although the problem considered there was different. Consequently, inequality (2.7) is equivalent to

$$\begin{aligned}&\left[(S_X^2-\sigma_\delta^2)^2 - \frac{t_{n-2(\gamma/2)}^2}{n}(\sigma_\delta^4 + \sigma_\delta^2 S_X^2)\right]\beta^2 - 2S_{XY}(S_X^2-\sigma_\delta^2)\beta \\ &\qquad + S_{XY}^2 - \frac{t_{n-2(\gamma/2)}^2}{n}\hat{\sigma}_\varepsilon^2 S_X^2 \le 0.\end{aligned} \tag{2.8}$$

It is worth noting that the confidence set in (2.8) is equal to the usual t-interval when $\sigma_\delta^2 = 0$ and its diameter can be infinite if the leading coefficient of the quadratic expression in β on the left of (2.8) is negative. Now the uniform convergence theorem of this confidence set is presented.

LEMMA 2.1. *Assume that model* (1.1)–(1.2) *holds with* σ_δ^2 *known. Then*

$$\lim_{n\to\infty} \sup_{\boldsymbol{\theta}\in\Omega} |P_{\boldsymbol{\theta}}(Z_n \le z) - \Phi(z)| = 0,$$

where

$$Z_n = \frac{\sqrt{n}[S_{XY} - \beta(S_X^2 - \sigma_\delta^2)]}{\sqrt{\beta^2(\sigma_\delta^4 + \sigma_\delta^2\sigma_X^2) + \sigma_\varepsilon^2\sigma_X^2}} \tag{2.9}$$

and $\Phi(z)$ *denotes the distribution function of* $N(0,1)$.

PROOF. Since subtracting m_u from u_i and α from Y_i in model (1.1)–(1.2) leaves Z_n unchanged, we can assume without loss of generality that $m_u = \alpha = 0$.

Write Z_n as $A_n + B_n$, where

$$A_n = \frac{\sqrt{n}\big[\sum X_i Y_i/n - \beta(\sum X_i^2/n - \sigma_\delta^2)\big]}{\sqrt{\beta^2(\sigma_\delta^4 + \sigma_\delta^2\sigma_X^2) + \sigma_\varepsilon^2\sigma_X^2}}, \qquad B_n = \frac{\sqrt{n}\bar{X}(\beta\bar{X} - \bar{Y})}{\sqrt{\beta^2(\sigma_\delta^4 + \sigma_\delta^2\sigma_X^2) + \sigma_\varepsilon^2\sigma_X^2}}.$$

To prove the result, it suffices to show that $A_n \to_L N(0,1)$ and $B_n \to_P 0$ uniformly in $\boldsymbol{\theta}$ as $n \to \infty$, where the notation $\to_L$ and $\to_P$ is used to denote convergence in distribution and in probability, respectively. By a direct computation, we have

$$\beta^2(\sigma_\delta^4 + \sigma_\delta^2\sigma_X^2) + \sigma_\varepsilon^2\sigma_X^2 \ge n^2\sigma_{\bar{X}}^2\sigma_{\beta\bar{\delta}-\bar{\varepsilon}}^2,$$

where $\bar{\delta} = \sum \delta_i/n$, $\bar{\varepsilon} = \sum \varepsilon_i/n$ and $\sigma_{\bar{X}}^2$ and $\sigma_{\beta\bar{\delta}-\bar{\varepsilon}}^2$ are the variances of $\bar{X}$ and $\beta\bar{\delta} - \bar{\varepsilon}$, respectively. Consequently,

$$|B_n| \le \frac{1}{\sqrt{n}}\left|\frac{\bar{X}}{\sigma_{\bar{X}}}\frac{(\beta\bar{\delta} - \bar{\varepsilon})}{\sigma_{\beta\bar{\delta}-\bar{\varepsilon}}}\right| \to 0$$

uniformly in $\boldsymbol{\theta}$ as $n \to \infty$.

Substituting $u_i + \delta_i$ and $\beta u_i + \varepsilon_i$ for X_i and Y_i, respectively, in A_n and by a straightforward manipulation, we have

$$A_n = \frac{1}{\sqrt{n}}\sum\left[\frac{u_i}{\sigma_u}\frac{\varepsilon_i}{\sigma_\varepsilon}\sigma_u\sigma_\varepsilon + \frac{\delta_i\varepsilon_i}{\sigma_\delta\sigma_\varepsilon}\sigma_\delta\sigma_\varepsilon - \left(\frac{\delta_i^2}{\sqrt{2}\sigma_\delta^2} - \frac{1}{\sqrt{2}}\right)\sqrt{2}\beta\sigma_\delta^2 - \frac{u_i\delta_i}{\sigma_u\sigma_\delta}\beta\sigma_u\sigma_\delta\right]$$
$$\times \frac{1}{\sqrt{\beta^2(\sigma_\delta^4 + \sigma_\delta^2\sigma_X^2) + \sigma_\varepsilon^2\sigma_X^2}},$$

which is denoted by $n^{-1/2}\sum_{i=1}^{n} T_i$. Note that the T_i's are independent, identically distributed random variables with mean 0 and variance 1. By the triangle inequality and the convexity of a cube function,

$$E|T_i|^3 \le E\left(\left|\frac{u_i\varepsilon_i}{\sigma_u\sigma_\varepsilon}\right| + \left|\frac{\delta_i\varepsilon_i}{\sigma_\delta\sigma_\varepsilon}\right| + \left|\frac{\delta_i^2}{\sqrt{2}\sigma_\delta^2} - \frac{1}{\sqrt{2}}\right| + \left|\frac{u_i\delta_i}{\sigma_u\sigma_\delta}\right|\right)^3$$
$$\le 16E\left\{\left|\frac{u_i\varepsilon_i}{\sigma_u\sigma_\varepsilon}\right|^3 + \left|\frac{\delta_i\varepsilon_i}{\sigma_\delta\sigma_\varepsilon}\right|^3 + \left|\frac{\delta_i^2}{\sqrt{2}\sigma_\delta^2} - \frac{1}{\sqrt{2}}\right|^3 + \left|\frac{u_i\delta_i}{\sigma_u\sigma_\delta}\right|^3\right\} \le K,$$

where K (> 0) is a constant which is independent of $\boldsymbol{\theta}$ since the distributions of u_i/σ_u, δ_i/σ_δ, $\varepsilon_i/\sigma_\varepsilon$ and $\delta_i^2/(\sqrt{2}\sigma_\delta^2)$ are the same. Now by the Berry–Esseen theorem [Chung (1974)]

$$\sup_z |P(A_n \le z) - \Phi(z)| \le \frac{3K}{\sqrt{n}}.$$

From this, it is easy to complete the proof. □

LEMMA 2.2. *Under the assumptions of Lemma* 2.1, $\forall\ \eta > 0$,

$$\lim_{n\to\infty} \sup_{\boldsymbol{\theta}\in\Omega} P_{\boldsymbol{\theta}}(|R_n - 1| > \eta) = 0,$$

where

$$R_n = \sqrt{\frac{\beta^2(\sigma_\delta^4 + \sigma_\delta^2\sigma_X^2) + \sigma_\varepsilon^2\sigma_X^2}{\beta^2(\sigma_\delta^4 + \sigma_\delta^2 S_X^2) + \hat{\sigma}_\varepsilon^2 S_X^2}}. \tag{2.10}$$

PROOF. It is easy to see that

$$\lim_{n\to\infty} P_{\boldsymbol{\theta}}(|R_n - 1| > \eta) = 0 \qquad \forall\ \boldsymbol{\theta} \in \Omega.$$

Our proof that this convergence is uniform in $\boldsymbol{\theta}$ is lengthy and hence is omitted. Interested readers may refer to Huwang (1991), Theorem 2.2.7. □

THEOREM 2.3. *Under model* (1.1)–(1.2) *with* σ_δ^2 *known*,

$$\lim_{n\to\infty} \sup_{\boldsymbol{\theta}\in\Omega} |P_{\boldsymbol{\theta}}(T_n^* \le t_{n-2(\gamma/2)}) - (1 - \gamma/2)| = 0, \tag{2.11}$$

where T_n^* *is given by* (2.4).

PROOF. Since $\Phi(t_{n-2(\gamma/2)}) \to 1 - \gamma/2$ as $n \to \infty$, by the triangle inequality it suffices to show that $T_n^* \to_L N(0, 1)$ uniformly in $\boldsymbol{\theta}$ as $n \to \infty$. By the fact that $T_n^* = Z_n R_n$, where Z_n and R_n are given by (2.9) and (2.10), respectively, the result follows easily from Lemmas 2.1 and 2.2. □

Note that Theorem 2.3 clearly implies that the two-sided confidence set (2.7) has coverage probability converging to $1 - \gamma$ uniformly in $\boldsymbol{\theta}$ and has the property of asymptotic honesty as defined in (1.3).

We now discuss a desirable property of confidence set (2.7). Let $\mathscr{C} = \{C(\mathbf{X}, \mathbf{Y});\ C(\mathbf{X}, \mathbf{Y}) \text{ satisfies } (1.3)\}$ be the set of all $100(1-\gamma)\%$ asymptotically honest confidence sets $C(\mathbf{X}, \mathbf{Y})$ for β. Then, on the basis of Gleser and Hwang's theorem, each confidence set in $\mathscr{C}$ will inevitably have a positive probability of having an infinite diameter. However, one would want this probability to be as small as possible. Therefore, among the class $\mathscr{C}$, it is desirable to seek a confidence set such that the supremum as well as the infimum probability over the parameter space of having an infinite diameter is asymptotically minimized.

THEOREM 2.4. *Assume that model* (1.1)–(1.2) *holds with* σ_δ^2 *known. Then, for each* $C(\mathbf{X}, \mathbf{Y}) \in \mathscr{C}$,

$$\limsup_{n\to\infty} \sup_{\boldsymbol{\theta}\in\Omega} P_{\boldsymbol{\theta}}[\textit{diameter of } C(\mathbf{X}, \mathbf{Y}) = \infty] \geq 1-\gamma. \tag{2.13}$$

PROOF. This result follows from the similar arguments of Gleser and Hwang (Theorem 1); see Huwang (1991), Lemma 2.3.2, for details.

In the following theorem we will establish that the proposed confidence set (2.7) minimizes the quantity on the left-hand side of (2.13), and we will find the value of

$$\limsup_{n\to\infty} \inf_{\boldsymbol{\theta}\in\Omega} P_{\boldsymbol{\theta}}[\text{diameter of } C(\mathbf{X}, \mathbf{Y}) = \infty].$$

THEOREM 2.5. *Assume that model* (1.1)–(1.2) *holds with* σ_δ^2 *known and that* σ_u^2 *varies as a function of* n. *As* $n \to \infty$,

$$P_{\sigma_u^2}[\textit{diameter of confidence set } (2.7) = \infty]$$

$$\to \begin{cases} 1-\gamma, & \textit{if } \sqrt{n}\sigma_u^2 \to 0, \\ \Phi(-c/(\sqrt{2}\sigma_\delta^2) + z_{\gamma/2}) - \Phi(-c/(\sqrt{2}\sigma_\delta^2) - z_{\gamma/2}), & \textit{if } \sqrt{n}\sigma_u^2 \to c, \\ 0, & \textit{if } \sqrt{n}\sigma_u^2 \to \infty, \end{cases} \tag{2.14}$$

where c *is any positive constant. As a consequence, confidence set* (2.7) *satisfies*

$$\lim_{n\to\infty} \sup_{\boldsymbol{\theta}\in\Omega} P_{\boldsymbol{\theta}}[\textit{diameter of confidence set } (2.7) = \infty] = 1-\gamma \tag{2.15}$$

and

$$\lim_{n\to\infty} \inf_{\boldsymbol{\theta}\in\Omega} P_{\boldsymbol{\theta}}[\textit{diameter of confidence set } (2.7) = \infty] = 0. \tag{2.16}$$

PROOF. First of all, note that the confidence set (2.7) has an infinite diameter if and only if the leading coefficient $A = (S_x^2 - \sigma_\delta^2)^2 - n^{-1} t_{n-2(\gamma/2)}^2 (\sigma_\delta^4 + \sigma_\delta^2 S_x^2)$ of the quadratic expression in β on the left of (2.8) is less than 0. By a straightforward manipulation, $A < 0$ is equivalent to

$$W^2 - (2nr + t_{n-2(\gamma/2)}^2 r)W + (n^2r^2 - nr^2 t_{n-2(\gamma/2)}^2) < 0, \tag{2.17}$$

where $\sigma_X^2 = \sigma_u^2 + \sigma_\delta^2$, $r = \sigma_\delta^2/\sigma_X^2$ and $W = nS_X^2/\sigma_X^2$ is a chi-square distribution with $n-1$ degrees of freedom. Since the leading coefficient of the quadratic expression in W on the left of (2.17) equals 1 and the discriminant is greater than 0,

$$\begin{aligned}
P_{\sigma_u^2}&(A<0)\\
&= P_{\sigma_u^2}\left\{W \in \frac{1}{2}\left[2nr + t_{n-2(\gamma/2)} r \left(t_{n-2(\gamma/2)} \pm \sqrt{8n + t^2_{n-2(\gamma/2)}}\right)\right]\right\}\\
&= P_{\sigma_u^2}\left\{\frac{W-(n-1)}{\sqrt{2(n-1)}}\right.\\
&\qquad\qquad \left. \in \frac{n(r-1)+1+t_{n-2(\gamma/2)} r\left[\frac{1}{2}t_{n-2(\gamma/2)} \pm \sqrt{2n + \frac{1}{4}t^2_{n-2(\gamma/2)}}\right]}{\sqrt{2(n-1)}}\right\},
\end{aligned}$$

which converges to the desired result by the fact that $[W-(n-1)]/\sqrt{2(n-1)} \to_L N(0,1)$ and $t_{n-2(\gamma/2)} \to z_{\gamma/2}$ as $n \to \infty$. □

REMARK 1. It is conjectured that every asymptotically honest confidence set for β would, under each of the three cases in (2.14), have a limiting probability of infinite length at least equal to the value in (2.14). Therefore, the confidence set (2.7) would be the best possible among $\mathscr{C}$. However, this is difficult to resolve.

REMARK 2. In addition to (2.16), confidence set (2.7) also has the desirable property that

$$\lim_{n\to\infty} P_{\boldsymbol{\theta}}[\text{diameter of confidence set (2.7)} = \infty] = 0 \qquad \forall\, \boldsymbol{\theta} \in \Omega,$$

and, in fact, this limit holds uniformly on any compact subset of Ω.

3. Confidence set for the slope and intercept. Based on the same idea used in constructing an asymptotically honest confidence set for β, a confidence set for (α, β) can be derived from the pivotal Q_n^* defined by

$$\begin{aligned}
Q_n^* &= n(\hat\alpha - \alpha, \hat\beta - \beta)\hat\Gamma^{*^{-1}}(\hat\alpha - \alpha, \hat\beta - \beta)'\\
&= \frac{n(\bar Y - \alpha - \beta\bar X)^2}{\beta^2\sigma_\delta^2 + \hat\sigma_\varepsilon^2} + \frac{n[S_{XY} - \beta(S_X^2 - \sigma_\delta^2)]^2}{\beta^2(\sigma_\delta^4 + \sigma_\delta^2 S_X^2) + \hat\sigma_\varepsilon^2 S_X^2}.
\end{aligned} \tag{3.1}$$

Here $\hat\Gamma^*$ is obtained from $\hat\Gamma$ in (2.2) by substituting $V^*(\hat\beta)$ of (2.6) and $S_v^{*^2} = \beta^2\sigma_\delta^2 + \hat\sigma_\varepsilon^2$ for $\hat V(\hat\beta)$ and S_v^2 in $\hat\Gamma$ and hence the parameter β is not estimated in $\hat\Gamma^*$. It is easy to see that Q_n^* converges in distribution to a chi-square with 2 degrees of freedom. In practice, it seems reasonable to approximate the distribution of $Q_n^*/2$ with that of F with 2 and $n-2$ degrees of freedom. Consequently, a $100(1-\gamma)\%$ confidence set for (α, β) is defined by

$$\tfrac{1}{2} Q_n^* < F_{2,\, n-2(\gamma)}, \tag{3.2}$$

where $F_{2,\, n-2(\gamma)}$ is the upper 100γ percentile of the F distribution with 2 and $n-2$ degrees of freedom.

Note that confidence set (3.2) is the usual F confidence region when $\sigma_\delta^2 = 0$. In fact, an explicit formula for (3.2) is obtainable which will be described below.

First of all, by taking $\alpha = \bar{Y} - \beta \bar{X}$ the minimum of Q_n^* over the space of α is achieved with the value

$$B_n(\beta) = \frac{n[S_{XY} - \beta(S_X^2 - \sigma_\delta^2)]^2}{\beta^2(\sigma_\delta^4 + \sigma_\delta^2 S_X^2) + \hat{\sigma}_\varepsilon^2 S_X^2}.$$

Therefore, the range of β, denoted by $C_s(\mathbf{X}, \mathbf{Y})$, can be first obtained by solving β from

$$\tfrac{1}{2} B_n(\beta) < F_{2,\, n-2(\gamma)},$$

or equivalently from

$$\begin{aligned} &\left[(S_X^2 - \sigma_\delta^2)^2 - \frac{2F_{2,n-2(\gamma)}}{n}(\sigma_\delta^4 + \sigma_\delta^2 S_X^2)\right]\beta^2 - 2(S_X^2 - \sigma_\delta^2) S_{XY}\beta \\ &\quad + S_{XY}^2 - \frac{2F_{2,n-2(\gamma)}}{n}\hat{\sigma}_\varepsilon^2 S_X^2 < 0. \end{aligned} \tag{3.3}$$

As a consequence,

$$C_s(\mathbf{X}, \mathbf{Y}) = \begin{cases} (r_1, r_2), & \text{if } L > 0, \\ (r_1, r_2)^c, & \text{if } L < 0 \text{ and } D > 0, \\ (-\infty, \infty), & \text{if } L < 0 \text{ and } D < 0, \end{cases}$$

where L, D, r_1 and r_2 are, respectively, the leading coefficient, the discriminant and the two roots of the quadratic expression in β on the left of (3.3). For each $\beta \in C_s(\mathbf{X}, \mathbf{Y})$, the range of α, denoted by $C_i(\mathbf{X}, \mathbf{Y})$, can be obtained by solving α from (3.2). After algebraic simplification,

$$C_i(\mathbf{X}, \mathbf{Y}) = \bar{Y} - \beta\bar{X} \pm \sqrt{[2F_{2,\, n-2(\gamma)} - B_n(\beta)](\beta^2\sigma_\delta^2 + \hat{\sigma}_\varepsilon^2)/n}.$$

In summary, this confidence set consists of (α, β) wherein $\beta \in C_s(\mathbf{X}, \mathbf{Y})$ and, for each β in $C_s(\mathbf{X}, \mathbf{Y})$, $\alpha \in C_i(\mathbf{X}, \mathbf{Y})$.

Confidence set (3.2) is asymptotically honest as expected and, in fact, it has the stronger property of uniform convergence which is established in the following theorem.

THEOREM 3.1. *Assume that model* (1.1)–(1.2) *holds with* σ_δ^2 *known. Then*

$$\lim_{n\to\infty} \sup_{\boldsymbol{\theta}\in\Omega} |\, P_{\boldsymbol{\theta}}(Q_n^* \le 2F_{2,\, n-2(\gamma)}) - (1-\gamma)\,| = 0.$$

PROOF. Since $P(\chi_2^2 \le 2F_{2,n-2(\gamma)}) \to 1-\gamma$, by the triangle inequality, it suffices to show that $Q_n^* \to_L \chi_2^2$ uniformly in $\boldsymbol{\theta}$ as $n \to \infty$. Rewrite Q_n^* as $N_1R_1 + N_2R_2$:

$$N_1 = \frac{n(\bar{Y}-\alpha-\beta\bar{X})^2}{\beta^2\sigma_\delta^2+\sigma_\varepsilon^2}, \qquad R_1 = \frac{\beta^2\sigma_\delta^2+\sigma_\varepsilon^2}{\beta^2\sigma_\delta^2+\hat{\sigma}_\varepsilon^2}, \qquad N_2 = Z_n^2, \qquad R_2 = R_n^2,$$

where Z_n and R_n are given by (2.9) and (2.10), respectively. Although N_1, R_1, N_2 and R_2 defined above depend on n, this is not made explicit in the notation. By Lemma 2.2 and a result similar to it, we have

$$\frac{\hat{\sigma}_\varepsilon^2-\sigma_\varepsilon^2}{\beta^2\sigma_\delta^2+\sigma_\varepsilon^2}(= R_1^{-1}-1) \to_P 0 \quad \text{and} \quad R_2^{1/2} \to_P 1$$

uniformly in $\boldsymbol{\theta}$ as $n \to \infty$, and hence R_1 and R_2 converge to 1 uniformly.

For each $\xi \in (0,1)$ and $t \in (0,\infty)$, it is easy to establish that

$$\begin{aligned} &P_{\boldsymbol{\theta}}\left(N_1+N_2 \le \frac{t}{1+\xi}\right) - P(\chi_2^2 \le t) - P_{\boldsymbol{\theta}}(R_1 \in [1\pm\xi]^c \text{ or } R_2 \in [1\pm\xi]^c) \\ &\qquad \le P_{\boldsymbol{\theta}}(N_1R_1+N_2R_2 \le t) - P(\chi_2^2 \le t) \end{aligned} \tag{3.4}$$

and

$$\begin{aligned} &P_{\boldsymbol{\theta}}\left(N_1+N_2 \le \frac{t}{1-\xi}\right) - P(\chi_2^2 \le t) + P_{\boldsymbol{\theta}}(R_1 \in [1\pm\xi]^c \text{ or } R_2 \in [1\pm\xi]^c) \\ &\qquad \ge P_{\boldsymbol{\theta}}(N_1R_1+N_2R_2 \le t) - P(\chi_2^2 \le t). \end{aligned} \tag{3.5}$$

Therefore, the absolute value of the right-hand side of (3.4) is bounded by the maximum of

$$\begin{aligned} &\left|P_{\boldsymbol{\theta}}\left(N_1+N_2 \le \frac{t}{1+\xi}\right) - P\left(\chi_2^2 \le \frac{t}{1+\xi}\right)\right| + \left|P\left(\chi_2^2 \le \frac{t}{1+\xi}\right) - P(\chi_2^2 \le t)\right| \\ &\qquad + P_{\boldsymbol{\theta}}(R_1 \in [1\pm\xi]^c \text{ or } R_2 \in [1\pm\xi]^c) \end{aligned}$$

and

$$\begin{aligned} &\left|P_{\boldsymbol{\theta}}\left(N_1+N_2 \le \frac{t}{1-\xi}\right) - P\left(\chi_2^2 \le \frac{t}{1-\xi}\right)\right| + \left|P\left(\chi_2^2 \le \frac{t}{1-\xi}\right) - P(\chi_2^2 \le t)\right| \\ &\qquad + P_{\boldsymbol{\theta}}(R_1 \in [1\pm\xi]^c \text{ or } R_2 \in [1\pm\xi]^c). \end{aligned}$$

Since $R_1 \to_P 1$ and $R_2 \to_P 1$ uniformly in $\boldsymbol{\theta}$ as $n \to \infty$ and ξ is arbitrary, the theorem will be proven if we can show that $N_1+N_2 \to_L \chi_2^2$ uniformly. Observe that N_1 has a χ_1^2 distribution because $\bar{Y}-\alpha-\beta\bar{X} \sim N(0, (\beta^2\sigma_\delta^2+\sigma_\varepsilon^2)/n)$. By Lemma 2.1, it follows that $N_2 \to_L \chi_1^2$ uniformly in $\boldsymbol{\theta}$ as $n \to \infty$. The result follows from the fact that N_1 and N_2 are independent. □

4. Confidence set for the intercept. The scenario for constructing an asymptotically honest confidence set for α is quite different from that of β described previously. From Γ given by (2.1), note that $\sqrt{n}(\hat{\alpha}-\alpha)$ has the asymptotic variance

$$\frac{m_u^2}{\sigma_u^4}[\beta^2(\sigma_\delta^4+\sigma_\delta^2\sigma_X^2)+\sigma_\varepsilon^2\sigma_X^2]+\beta^2\sigma_\delta^2+\sigma_\varepsilon^2, \tag{4.1}$$

which depends on β but not α. If we estimate all parameters (including β) in (4.1) and apply the traditional method to construct a confidence set for α, the resultant one will have an almost surely finite diameter and hence have a zero confidence level due to the results of Gleser and Hwang. Therefore, the approach to the asymptotically honest confidence set for β used previously cannot be adapted in this situation. Scheffé's projection method is another possible way to attack the problem. A $100(1-\gamma)\%$ confidence set for α can be obtained by projecting (3.2) onto the α-axis. Obviously, the projected confidence set is asymptotically honest. However, this approach needs tedious numerical computations and the result may be too conservative. Hence it will not be pursued here.

Acknowledgment. I would like to thank my thesis advisor, J. T. Gene Hwang, for his guidance in this work. This paper is part of the author's dissertation presented to Cornell University in partial fulfillment of the requirements for a Ph.D. degree.

REFERENCES

ANDERSON, T. W. (1984). Estimating linear statistical relationships. *Ann. Statist.* **12** 1–45.

CHUNG, K. L. (1974). *A Course in Probability Theory*, 2nd ed. Academic Press, New York.

CREASY, M. A. (1956). Confidence limits for the gradient in the linear functional relationship. *J. Roy. Statist. Soc. Ser. B* **18** 65–69.

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

GLESER, L. J. and HWANG, J. T. (1987). The nonexistence of $100(1-\alpha)\%$ confidence sets of finite expected diameter in errors-in-variables and related models. *Ann. Statist.* **15** 1351–1362.

HUWANG, L. (1991). Good coverage probability confidence sets in linear errors-in-variables models. Ph.D. dissertation, Statistics Center, Cornell Univ.

HWANG, J. T. (1995). Fieller's problems and resampling techniques. *Statistica Sinica* **5** 161–171.

KENDALL, M. G. and STUART, A. (1979). *The Advanced Theory of Statistics* **2**, 4th ed. Hafner, New York.

LI, K. C. (1989). Honest confidence regions for nonparametric regression. *Ann. Statist.* **17** 1001–1008.

MADANSKY, A. (1959). The fitting of straight lines when both variables are subject to error. *J. Amer. Statist. Assoc.* **54** 173–205.

MORAN, P. A. P. (1971). Estimating structural and functional relationships. *J. Multivariate Anal.* **1** 232–255.

REIERSOL, O. (1950). Identifiability of a linear relation between variables which are subject to error. *Econometrica* **18** 375–389.

NATIONAL TSING HUA UNIVERSITY
INSTITUTE OF STATISTICS
HSINCHU, TAIWAN