Geometric View of Measurement Errors

DIARMUID O'DRISCOLL AND DONALD E. RAMIREZ

Abstract

The slope of the best fit line from minimizing the sum of the squared oblique errors is the root of a polynomial of degree four. This geometric view of measurement errors is used to give insight into the performance of various slope estimators for the measurement error model including an adjusted fourth moment estimator introduced by Gillard and Iles (2005) to remove the jump discontinuity in the estimator of Copas (1972). The polynomial of degree four is associated with a minimun deviation estimator. A simulation study compares these estimators showing improvement in bias and mean squared error.

Department of Mathematics and Computer Studies, Mary Immaculate College, Limerick, Ireland *E-mail address: diarmuid.odriscoll@mic.ul.ie*

Department of Mathematics, University of Virginia, Charlottesville, VA 22904 *E-mail address: der@virginia.edu*

Keywords Oblique errors; Measurement errors; Maximum likelihood estimation; Moment estimation.

Mathematics Subject Classification 62J05; 62G05.

1 Introduction

With ordinary least squares OLS($y|x$) regression, we have data $\{(x_1, Y_1|X = x_1), ..., (x_n, Y_n|X = x_n)\}$ and we minimize the sum of the squared *vertical errors* to find the *best-fit* line $y = h(x) = \beta_0 + \beta_1 x$. With OLS($y|x$) it is assumed that the independent or causal variable is measured without error. The measurement error model has wide interest with many applications. See for example Carroll *et al.* (2006) and Fuller (1987). The comparison of measurements by two analytical methods in clinical chemistry is often based on regression analysis. There is no causal or independent variable in this type of analysis. The most frequently used method to determine any systematic difference between two analytical methods is OLS($y|x$) which has several shortcomings when both measurement sets are subject to error. Linnet(1993) states that "it is rare that one of the (measurement) methods is without error." Linnet(1999) further states that " A systematic difference between two (measurement) methods is identified if the estimated intercept differs significantly from zero (constant difference) or if the slope deviates significantly from 1 (proportional difference)." Our paper concentrates on how to determine whether or not there is a significant difference between two measurement instruments using a Monte Carlo simulation; that is, we concentrate our studies about a true slope of 1. As in the regression procedure of Deming (1943), to account for both sets of errors, we determine a fit so that both the squared vertical and the squared horizontal errors will be minimized. The squared vertical errors are the squared distances from (x, y) to $(x, h(x))$ and the squared horizontal errors are the squared distances from (x, y) to $(h^{-1}(y), y)$. As a compromise, we will consider oblique errors. All of the estimated regression models we consider (including the geometric mean and perpendicular methods) are contained in the parametrization (with $0 \le \lambda \le 1$) of the line from $(x, h(x))$ to $(h^{-1}(y), y)$.

We review the Oblique Error Method in Section 2. In Section 3, we review the geometric mean and perpendicular error models. In Section 4, we show how the geometric mean slope is a natural estimator for the slope in the measurement error (error-in-variables) model. Section 5 shows a relationship between the maximum likelihood estimator in the measurement error model and the geometric mean estimator. We give a case study to illustrate the effects that erroneous assumptions for the ratio of variance of errors can have on the maximum likelihood estimators. Section 6 discusses a fourth moment estimator and shows a circular relationship to the maximum likelihood estimator. Section 7 develops a minimum deviation estimator derived by minimizing Equation (2) in Section 2 with respect to λ for fixed β_1. Section 8 contains our Monte Carlo simulations where we compare these estimators. Supporting Maple worksheets are available from the link http://people.virginia.edu/˜der/ODriscoll_Ramirez/.

2 Minimizing Squared Oblique Errors

¿From the data point (x_i, y_i) to the fitted line $y = h(x) = \beta_0 + \beta_1 x$ the vertical length is $a_i = |y_i - \beta_0 - \beta_1 x_i|$, the horizontal length is $b_i = |x_i - (y_i - \beta_0)/\beta_1| = |(\beta_1 x_i - y_i + \beta_0)/\beta_1| = |a_i/\beta_1|$ and the perpendicular length is $h_i = a_i/\sqrt{1+\beta_i^2}$. With standard notation, $S_{xx} = \sum_{i=1}^n (x_i - \overline{x})^2$, $S_{yy} = \sum_{i=1}^n (y_i - \overline{y})^2$, $S_{xy} = \sum_{i=1}^n (x_i - \overline{x})(y_i - \overline{y})$ with the correlation $\rho = S_{xy}/\sqrt{S_{xx}S_{yy}}$.

For the oblique length from (x_i, y_i) to $(h^{-1}(y_i) + \lambda(x_i - h^{-1}(y_i)), y_i + \lambda(h(x_i) - y_i))$, the horizontal error is $(1-\lambda)b_i = (1-\lambda)a_i/|\beta_1|$ and the vertical error is λa_i. The sum of squared horizontal, respectively vertical, errors are given by $SSE_h(\beta_0, \beta_1, \lambda) = \left(\sum_{i=1}^n a_i^2\right)/\beta_1^2$ and $SSE_v(\beta_0, \beta_1, \lambda) = \sum_{i=1}^n a_i^2$. In a comprehensive paper by Riggs *et al.* (1978), the authors place great emphasis on the importance of equations being dimensionally correct, since it is from these equations that the slope estimators are derived. In particular the authors state that: "It is a poor method indeed whose results depend upon the particular units chosen for measuring the variables ... and that invariance under linear transformations is equivalent to requiring the method to be dimensionally correct." So that our equation is dimensionally correct we consider

$$SSE_o(\beta_0, \beta_1, \lambda) = (1-\lambda)^2 \frac{SSE_h}{\widetilde{\sigma}_\delta^2} + \lambda^2 \frac{SSE_v}{\widetilde{\sigma}_\tau^2} \tag{1}$$

where $\{\widetilde{\sigma}_\delta^2, \widetilde{\sigma}_\tau^2\}$ are Madansky's moment estimators of the variance in the horizontal, respectively vertical, directions. In Section 4, we show that this is equivalent to using

$$SSE_o(\beta_0, \beta_1, \lambda) = (1-\lambda)^2 S_{yy} SSE_h + \lambda^2 S_{xx} SSE_v. \tag{2}$$

Similar to that shown in O'Driscoll, Ramirez and Schmitz (2008), the solution of $\partial SSE_o/\partial\beta_0 = 0$ is given by $\beta_0 = \overline{y} - \beta_1 \overline{x}$ and hence

$$SSE_o(\beta_0, \beta_1, \lambda) = \left((1-\lambda)^2 S_{yy}/\beta_1^2 + \lambda^2 S_{xx}\right)(S_{yy} - 2\beta_1 S_{xy} + \beta_1^2 S_{xx}).$$

The solutions of $\partial SSE_o/\partial\beta_1 = 0$ are then the roots of the fourth degree polynomial equation in β_1, namely

$$P_4(\beta_1) = \lambda^2 \sqrt{\frac{S_{xx}}{S_{yy}}} \frac{S_{xx}}{S_{yy}} \beta_1^4 - \lambda^2 \frac{S_{xx}}{S_{yy}} \rho \beta_1^3 + (1-\lambda)^2 \rho \beta_1 - (1-\lambda)^2 \sqrt{\frac{S_{yy}}{S_{xx}}} = 0. \tag{3}$$

With $\lambda = 1$ we recover the minimum squared vertical errors with estimated slope β_1^{ver}, and with $\lambda = 0$ we recover the minimum squared horizontal errors with estimated slope β_1^{hor}.

For each fixed $\lambda \in [0.1]$, there corresponds $\beta_1 \in [\beta_1^{ver}, \beta_1^{hor}]$ which satisfies Equation (3), and conversely, for each fixed $\beta_1 \in [\beta_1^{ver}, \beta_1^{hor}]$, there corresponds $\lambda \in [0.1]$ such that minimizing the sum of the squared oblique errors has estimated slope β_1. We measure the angle θ_λ of the oblique projection associated with λ using the line segments (x, y) to $(x, h(x))$ and $(x, h(x))$ to $(h^{-1}(y), y)$. When the slope β_1 is close to one, for

λ near one we anticipate θ_λ to be near $45°$ and for λ is close to zero we anticipate θ_λ to be near $135°$. The angles are computed from the Law of Cosines.

A similar argument to that of O'Driscoll *et al.* (2008) shows that $P_4(\beta_1)$ has exactly two real roots, one positive and one negative with the global minimum being the positive (respectively negative) root corresponding to the sign of S_{xy}. Riggs *et al.* (1978) in Equation (119) also noted the role of the roots of a similar quartic equation in determining the slope estimators.

3 Minimizing Squared Perpendicular and Squared Geometric Mean Errors

The perpendicular error model dates back to Adcock (1878) who introduced it as a procedure for fitting a straight line model to data with error measured in both the x and y directions. For squared perpendicular errors Adcock minimized $SSE_{per}(\beta_0, \beta_1) = \sum_{i=1}^{n} a_i^2/(1+\beta_1^2)$ with solutions $\beta_0^{per} = \overline{y} - \beta_1^{per}\overline{x}$ and

$$\beta_1^{per} = \frac{(S_{yy} - S_{xx}) \pm \sqrt{(S_{yy} - S_{xx})^2 + 4S_{xy}^2}}{2S_{xy}}, \tag{4}$$

provided $S_{xy} \neq 0$. However, in this case, the equation which minimizes $SSE_{per}(\beta_0, \beta_1)$ is dimensionally incorrect unless x and y are measured in the same units.

For squared geometric mean errors, we minimize $SSE_{gm}(\beta_0, \beta_1) = \sum_{i=1}^{n} \left(\sqrt{|a_i b_i|}\right)^2 = \sum_{i=1}^{n} a_i^2/|\beta_1|$ with solutions $\beta_0^{gm} = \overline{y} - \beta_1^{gm}\overline{x}$ and

$$\beta_1^{gm} = \pm\sqrt{S_{yy}/S_{xx}}. \tag{5}$$

Proposition 1. *The geometric mean estimator has oblique parameter $\lambda = 1/2$.*

Proof: For $\beta_1^{gm} = \sqrt{S_{yy}/S_{xx}}$,we solve the quadratic equation $P_4(\beta_1^{gm}) = 0$ for λ. This equation reduces to a linear equation whose root is $\lambda = 1/2.\square$

4 Measurement Error Model and Second Moment Estimation

We now consider the measurement error (errors-in-variables) model as follows. In this paper it is assumed that X and Y are random variables with respective finite variances σ_X^2 and σ_Y^2, finite fourth moments and have the linear functional relationship $Y = \beta_0 + \beta_1 X$. The observed data $\{(x_i, y_i), 1 \leq i \leq n\}$ are subject to error by $x_i = X_i + \delta_i$ and $y_i = Y_i + \tau_i$ where it is also assumed that δ is $N(0, \sigma_\delta^2)$ and τ is $N(0, \sigma_\tau^2)$. In our simulation studies we will use an exponential distribution for X.

It is well known, in a measurement error model, that the expected value for β_1^{ver} is attenuated to zero by the attenuating factor $\sigma_X^2/(\sigma_\delta^2 + \sigma_X^2)$, called the reliability ratio by Fuller (1987). Similarly the expected value for β_1^{hor} is amplified to infinity by the amplifying factor $(\sigma_Y^2 + \sigma_\tau^2)/\sigma_Y^2$. Thus for the measurement error model, when both the vertical and horizontal models are reasonable, a compromise estimator such as the geometric mean estimator β_1^{gm} is hoped to have improved efficiency. However, Lindley and El-Sayyad (1968) proved that the expected value of β_1^{gm} is biased unless $\sigma_\tau^2/\sigma_Y^2 = \sigma_\delta^2/\sigma_X^2$.

Madansky's moment estimators for $\{\sigma_\delta^2, \sigma_\tau^2\}$ are

$$\begin{aligned} \widetilde{\sigma}_\delta^2 &= \frac{S_{xx}}{n} - \frac{S_{xy}}{n\beta_1}, \\ \widetilde{\sigma}_\tau^2 &= \frac{S_{yy}}{n} - \frac{\beta_1 S_{xy}}{n}. \end{aligned} \tag{6}$$

If σ_δ is known or can be approximated, Madansky used the first of the equations in Equation (6) to derive an estimator for β_1 but Riggs *et al.* (1978) in Figure 5 produced an example for which this estimator performs poorly. In general σ_δ is not known and concentration focuses on estimating the true error ratio $\sigma_\tau^2/\sigma_\delta^2$. Our Monte Carlo simulation illustrates how poor estimates for the error ratio may lead to large biases using the MLE estimator. It would be interesting to determine if there is a slope estimator β_1 that is a fixed point for $\beta_1 = \widetilde{\sigma}_\tau(\beta_1)/\widetilde{\sigma}_\delta(\beta_1)$. This can be achieved with the geometric mean estimator β_1^{gm}.

Proposition 2. β_1^{gm} *is a fixed point of the ratio function* $\beta_1 = \dfrac{\widetilde{\sigma}_\tau(\beta_1)}{\widetilde{\sigma}_\delta(\beta_1)}$.

Proof: Rewrite $n\widetilde{\sigma}_\delta^2 = S_{xx} - S_{xy}\beta_1^{gm} = S_{xx} - S_{xy}\sqrt{S_{xx}/S_{yy}}$ and $n\widetilde{\sigma}_\tau^2 = S_{yy} - S_{xy}\beta_1^{gm} = S_{yy} - S_{xy}\sqrt{S_{yy}/S_{xx}}$, from which $\widetilde{\sigma}_\tau/\widetilde{\sigma}_\delta = \sqrt{S_{yy}/S_{xx}}$.□

We return to the assertion made in Section 1. A natural standardized weighed average for the oblique model is shown in Equation (1) and using the fixed point solution of Proposition 2 in this equation yields the equivalent model given in Equation (2).

5 The Maximum Likelihood Estimator

If the ratio of the error variances $\kappa = \sigma_\tau^2/\sigma_\delta^2$ is assumed finite, then Madansky (1959), among others, showed that the maximum likelihood estimator for the slope is

$$\beta_1^{mle} = \frac{(S_{yy} - \kappa S_{xx}) + \sqrt{(S_{yy} - \kappa S_{xx})^2 + 4\kappa\rho^2 S_{xx}S_{yy}}}{2\rho\sqrt{S_{xx}S_{yy}}} \tag{7}$$

It also follows that if $\kappa = 1$ in Equation (7) then the MLE (often called the Deming Regression estimator) is equivalent to the perpendicular estimator, β_1^{per}. Conversely, if the MLE is β_1^{per} then $\kappa = 1$. In the particular case where $S_{xx} = S_{yy}$ then β_1^{per} has a λ value of 0.5. We note that S_{yy}/S_{xx} is a good estimator of $\sigma_y^2/\ \sigma_x^2$, but in general, it is not a good estimator of the error ratio $\kappa = \sigma_\tau^2/\ \sigma_\delta^2$. In Section 6, we discuss a moment estimator $\widetilde{\kappa}$ for κ.

The MLE is a function of $\{\rho, \kappa, S_{xx}/S_{yy}\}$. Table 1 gives the corresponding β_1^{mle} value for typical values. For fixed $\{\kappa, \rho\}$, the values for β_1^{mle} in each column of Table 1 decrease. As expected, with $\kappa = 1$ and $S_{xx} = S_{yy}$, the maximum likelihood estimator agrees with the geometric mean estimator, both being equal to 1.00.

Table 1

Values for β_1^{mle} for typical $\{\rho, \kappa, S_{xx}/S_{yy}\}$

$\kappa =$	0.500	0.500	0.500	0.500	1.000	1.000	1.000	1.000	2.000	2.000	2.000	2.000
$\rho =$	0.200	0.400	0.600	0.800	0.200	0.400	0.600	0.800	0.200	0.400	0.600	0.800
$S_{xx}/S_{yy} = 1/2$	5.396	2.828	2.016	1.632	3.799	2.219	1.750	1.535	1.414	1.414	1.414	1.414
$S_{xx}/S_{yy} = 1$	2.686	1.569	1.237	1.086	1.000	1.000	1.000	1.000	0.372	0.638	0.808	0.921
$S_{xx}/S_{yy} = 2$	0.707	0.707	0.707	0.707	0.263	0.451	0.571	0.651	0.185	0.354	0.496	0.613

In Table 2 we record the corresponding obliqueness parameter λ for the maximum likelihood model for these typical values. Small values near 0 support OLS($x|y$), denoted by β_1^{hor}, and large values near 1 support OLS($y|x$), denoted by β_1^{ver}. For fixed $\{\kappa, \rho\}$, the values for the obliqueness parameter λ in each column of Table 2 increase indicating the model moves from β_1^{hor} towards β_1^{ver}. With $\kappa = S_{yy}/S_{xx}$, $\beta_1^{mle} = \beta_1^{gm}$ as shown by the cells of Table 2 with $\lambda = 0.500$.

Table 2
Values for λ for typical $\{\rho, \kappa, S_{xx}/S_{yy}\}$

$\kappa =$	0.500	0.500	0.500	0.500	1.000	1.000	1.000	1.000	2.000	2.000	2.000	2.000
$\rho =$	0.200	0.400	0.600	0.800	0.200	0.400	0.600	0.800	0.200	0.400	0.600	0.800
$S_{xx}/S_{yy} = 1/2$	0.033	0.111	0.197	0.273	0.089	0.223	0.316	0.375	0.500	0.500	0.500	0.500
$S_{xx}/S_{yy} = 1$	0.089	0.223	0.316	0.375	0.500	0.500	0.500	0.500	0.911	0.777	0.684	0.625
$S_{xx}/S_{yy} = 2$	0.500	0.500	0.500	0.500	0.911	0.776	0.684	0.625	0.967	0.889	0.803	0.727

The Madansky's moment estimators $\{\widetilde{\sigma}_\delta^2, \widetilde{\sigma}_\tau^2\}$ depend on the choice of β_1. In Table 3, we record the effect of varying slopes on the moments and their ratio when computable.

Table 3
Error ratios for Madansky's moment estimators for varying β_1

	$\widetilde{\sigma}_\delta^2$	$\widetilde{\sigma}_\tau^2$	$\frac{\widetilde{\sigma}_\tau^2}{\widetilde{\sigma}_\delta^2}$
β_1^{ver}	0	$\frac{1-\rho^2}{n} S_{yy}$	∞
β_1^{hor}	$\frac{1-\rho^2}{n} S_{xx}$	0	0
β_1^{gm}	$\frac{1-\rho}{n} S_{xx}$	$\frac{1-\rho}{n} S_{yy}$	$\frac{S_{yy}}{S_{xx}}$
β_1^{per}	$\frac{1}{2} \frac{S_{xx}+S_{yy}-\sqrt{(S_{xx}-S_{yy})^2+4\rho^2 S_{xx} S_{yy}}}{n}$	$\frac{1}{2} \frac{S_{xx}+S_{yy}-\sqrt{(S_{xx}-S_{yy})^2+4\rho^2 S_{xx} S_{yy}}}{n}$	1
β_1^{mle}	$\frac{1}{2} \frac{S_{xx}+\frac{S_{yy}}{\kappa}-\sqrt{(S_{xx}-\frac{S_{yy}}{\kappa})^2+4\rho^2 S_{xx} \frac{S_{yy}}{\kappa}}}{n}$	$\frac{1}{2} \frac{\kappa S_{xx}+S_{yy}-\sqrt{(\kappa S_{xx}-S_{yy})^2+4\rho^2 \kappa S_{xx} S_{yy}}}{n}$	κ

In the next section, we introduce a second moment estimator for κ and a fourth moment estimator for β_1.

6 Fourth Moment Estimation

When κ is unknown, Solari (1969) showed that the maximum likelihood estimator for the slope β_1 does not exist, as the maximum likelihood surface has a saddle point at the critical value. Earlier Lindley and El-Sayyad (1968) suggested, in this case, that the maximum likelihood method fails as the estimator would be the geometric mean estimator which converges to the wrong value. Sprent (1970) pointed out the result of Solari does not imply that the maximum likelihood principle has failed, but rather that the likelihood surface has no maximum value at the critical value.

Copas (1972) offered some advice for using the maximum likelihood method. He assumed the data has rounding-off errors in the observations which allows for an approximated likelihood function to be used, and that this approximated likelihood function is bounded. His estimator for the slope has the rule

$$\beta_1^{cop} = \begin{cases} \beta_1^{ver} & \text{if } \sum y_i^2 < \sum x_i^2 \\ \beta_1^{hor} & \text{if } \sum y_i^2 > \sum x_i^2 \end{cases} ,$$

so the ordinary least squares estimators are used depending on the whether $|\beta_1^{gm}| < 1$ or $|\beta_1^{gm}| > 1$.

The Copas estimator is *not* continuous in the data as a small change in data can switch the direction of the inequality $\sum y_i^2 < \sum x_i^2$ which will cause a jump discontinuity in the estimator β_1^{cop}. To achieve continuity in the data, we adjust the range of the fourth moment estimator β_1^{mom} described in Gillard and Iles (2005) to account for admissible values for $\{\sigma_\delta^2, \sigma_\tau^2\}$. See also Gillard and Iles (2010).

The basic second moment estimators for $\widetilde{\sigma}_\delta^2$ and $\widetilde{\sigma}_\tau^2$ are shown in Equation (6). Since variances must be positive, we have the admissible range for the moment estimator for $\widetilde{\beta}_1$ as

$$\beta_1^{ver} = \frac{S_{xy}}{S_{xx}} < \widetilde{\beta}_1 < \frac{S_{yy}}{S_{xy}} = \beta_1^{hor}. \tag{8}$$

Set $S_{xxxy} = \sum(x_i - \overline{x})^3(y_i - \overline{y})$ and similarly for S_{xyyy}. Following Gillard and Iles (2005), from Equations (22) and (24)), the fourth moment equations of interest are

$$\begin{aligned} \frac{S_{xxxy}}{n} &= \widetilde{\beta}\widetilde{\mu}_4 + 3\widetilde{\beta}\widetilde{\sigma}^2\widetilde{\sigma}_\delta^2 \\ \frac{S_{xyyy}}{n} &= \widetilde{\beta}^3\widetilde{\mu}_4 + 3\widetilde{\beta}\widetilde{\sigma}^2\widetilde{\sigma}_\varepsilon^2 \end{aligned} \tag{9}$$

with $\widetilde{\mu}_4$ denoting the fourth central moment for the underlying distribution of X. The four equations from (6) and (9) allow for a moment solution for β_1 as

$$\widetilde{\beta}_1 = \sqrt{\frac{S_{xyyy} - 3S_{xy}S_{yy}}{S_{xxxy} - 3S_{xy}S_{xx}}}. \tag{10}$$

In our simulation study, $\widetilde{\beta}_1$ was well-defined around 99% of the time. If the radicand is negative, we recommend using the geometric mean estimator.

To satisfy Equation (8) we define β_1^{mom} as

$$\beta_1^{mom} = \begin{cases} \beta_1^{ver} & \text{if } \widetilde{\beta}_1 \leq \beta_1^{ver} \\ \widetilde{\beta}_1 & \text{if } \beta_1^{ver} \leq \widetilde{\beta}_1 \leq \beta_1^{hor} \\ \beta_1^{hor} & \text{if } \widetilde{\beta}_1 \geq \beta_1^{hor} \end{cases}. \tag{11}$$

This is a Copas-type estimator with the moment estimator $\widetilde{\beta}_1$ used to "smooth" out the jump discontinuity inherent in the Copas estimator. We next study the circular relationship between this moment estimator and the maximum likelihood estimator with fixed κ.

We will define the moment estimator $\kappa(\beta_1)$ as a function of β_1, then use this value to compute $\beta_1^{mle}(\kappa)$ as a function of κ. Finally, we note that $\beta_1^{mle}(\kappa(\widetilde{\beta}_1)) = \widetilde{\beta}_1$, showing the circular relationship between the estimators $\{\widetilde{\beta}_1, \beta_1^{mle}\}$. Thus our moment estimator also has the functional form of the maximum likelihood estimator with fixed κ.

Set $\widetilde{\kappa}(\widetilde{\beta}_1) = \widetilde{\sigma}_\tau^2/\widetilde{\sigma}_\delta^2$ so

$$\widetilde{\kappa}(\widetilde{\beta}_1) = \frac{S_{yy} - \widetilde{\beta}_1\rho\sqrt{S_{xx}S_{yy}}}{S_{xx} - \rho/\widetilde{\beta}_1\sqrt{S_{xx}S_{yy}}}. \tag{12}$$

We use $\widetilde{\kappa}(\widetilde{\beta}_1)$ in Equation (7) to determine $\beta_1^{mle}(\widetilde{\kappa}(\widetilde{\beta}_1))$. As $\widetilde{\beta}_1 \to \beta_1^{hor}$ the numerator in Equation (12) tends to zero so $\widetilde{\kappa}(\widetilde{\beta}_1) \to 0$ and $\beta_1^{mle}(\widetilde{\kappa}(\widetilde{\beta}_1)) \to \beta_1^{hor}$; similarly as $\widetilde{\beta}_1 \to \beta_1^{ver}$ the denominator in Equation (12) tends to zero so $\widetilde{\kappa}(\widetilde{\beta}_1) \to \infty$ and $\beta_1^{mle}(\widetilde{\kappa}(\widetilde{\beta}_1)) \to \beta_1^{ver}$. A stronger result is given in the following Proposition.

Proposition 3. *For each β_1, $\beta_1^{mle}(\widetilde{\kappa}(\beta_1)) = \beta_1$ and in particular $\beta_1^{mle}(\widetilde{\kappa}(\widetilde{\beta}_1)) = \widetilde{\beta}_1$.*

Proof: In Equation (7) solve $\beta_1^{mle}(\kappa) = \beta_1$ for $\kappa = \kappa_0$, and then check that κ_0 is the same as in Equation (12).□

An example helps to demonstrate the smoothing achieved with the moment estimator β_1^{mom}. Assume $\{\rho = 0.5, S_{xx} = 1, S_{xxxy} = 10, S_{xyyy} = 5\}$. Equation (8) requires that $0.13029 \leq S_{yy} \leq 1.31862$. As S_{yy} varies over the admissible values for S_{yy}, $\widetilde{\kappa}(\widetilde{\beta}_1)$ varies over $[0, \infty]$ and $\widetilde{\beta}_1$ varies over $[\beta_1^{ver}, \beta_1^{hor}]$ and $\beta_1^{mle}(\widetilde{\kappa}(\widetilde{\beta}_1)) = \widetilde{\beta}_1$, a surprising result.

Table 4
Slope Estimates with $\{\rho = 0.5, S_{xx} = 1, S_{xxxy} = 10, S_{xyyy} = 5\}$

S_{yy}	β_1^{ver}	$\widetilde{\beta}_1$	β_1^{hor}	$\widetilde{\kappa}(\widetilde{\beta}_1)$	β_1^{mle}
0.1303	0.1805	0.7219	0.7219	0.0000	0.7219
0.2000	0.2236	0.7222	0.8944	0.0558	0.7222
0.4000	0.3164	0.7145	1.2649	0.3123	0.7145
0.6000	0.3873	0.6977	1.5492	0.7412	0.6977
0.8000	0.4472	0.6734	1.7889	1.4850	0.6734
1.0000	0.5000	0.6417	2.0000	3.0760	0.6417
1.2000	0.5477	0.6020	2.1909	9.6582	0.6020
1.3186	0.5742	0.5742	2.2966	∞	0.5741

7 Minimum Deviation Estimation

¿From Section 1 with fixed β_1 the solution of $\partial SSE_o/\partial\lambda = 0$ is given by $\lambda = S_{yy}/(S_{yy}+\beta_1^2 S_{xx})$. Substituting β_1^{mom} for β_1 in this result for λ produces a Minimum Deviation type estimator which we denote by β_1^{md}, with $\beta_1^{ver} \leq \beta_1^{md} \leq \beta_1^{hor}$. Our simulation studies will support the efficiency of this Minimum Deviation estimator.

Riggs *et al.* (1978) observes for $S_{yy}/S_{xx} = 1$ that the geometric mean estimator is unbiased when $\kappa = 1$, negatively biased when $\kappa < 1$, and positively biased when $\kappa > 1$. The authors also state that "no one method of estimating β_1 is the best method under all circumstances." To determine the efficiency of these estimators we conduct a Monte Carlo simulation in the next section.

8 Monte Carlo Simulation

Our Monte Carlo simulation uses X with an exponential distribution with mean $\mu_X = 10$ (and $\sigma_X = 10$) and $Y = X$ so $\beta_1 = 1$ and $\beta_0 = 0$. Both X and Y are subject to errors σ_δ^2, respectively σ_τ^2 where $(\sigma_\delta^2,\ \sigma_\tau^2) \in \{1,4,9\}\times\{1,4,9\}$. The sample size n is chosen as 100.

The first simulation, with the number of replications $R = 100$, summarized in Table 5, reports on the bias in the MLE estimator in using a misspecified value of κ. For $(\sigma_\delta^2,\ \sigma_\tau^2) \in \{1,4,9\}\times\{1,4,9\}$, κ ranges with ratios from 1 : 1 to 9 : 1 The true error ratios of κ are recorded in the first row and the assumed error ratios $\kappa^\#$ which are used to compute β_1^{mle} are recorded in the first column, both in ascending order.

Table 5
Percentage Bias of MLE estimator for the assumed ratios $\kappa^\#$ for varying values of $\kappa = \sigma_\tau^2/\sigma_\delta^2$
$\{\beta_1 = 1, \beta_0 = 0, n = 100, R = 100\}$

$\{\kappa^\#, \kappa\}$	1 : 9	1 : 4	4 : 9	1 : 1	4 : 4	9 : 9	9 : 4	4 : 1	9 : 1
1 : 9	0.166	0.502	2.164	0.870	3.663	7.995	8.723	3.592	9.282
1 : 4	−0.914	−0.012	0.811	0.666	2.807	6.087	7.351	3.067	8.265
4 : 9	−2.066	−0.564	−0.643	0.445	1.878	3.999	5.838	2.496	7.137
1 : 1	−4.067	−1.541	−3.184	0.051	0.218	0.266	3.083	1.467	5.058
4 : 4	−4.067	−1.541	−3.184	0.051	0.218	0.266	3.083	1.467	5.058
9 : 9	−4.067	−1.541	−3.184	0.051	0.218	0.266	3.083	1.467	5.058
9 : 4	−5.957	−2.495	−5.590	−0.342	−1.417	−3.330	0.338	0.437	2.936
4 : 1	−6.956	−3.016	−6.856	−0.561	−2.310	−5.230	−1.161	−0.136	1.748
9 : 1	−7.840	−3.489	−7.973	−0.763	−3.119	−6.899	−2.513	−0.663	0.657

As expected, the values for $\widetilde{\kappa} = \kappa$ show the smallest bias, and in each column for a given κ the bias shows that the estimated slope moves from over estimating the true value to under estimating the true value of $\beta_1 = 1$. This was anticipated since for $\widetilde{\kappa}$ near zero the maximum likelihood estimator favors β_1^{hor} which over estimates $\beta_{1;}$ and correspondingly, for $\widetilde{\kappa}$ near one the maximum likelihood estimator favors β_1^{ver} which under estimates β.

We conducted a second large scale Monte Carlo simulation study with $R = 1000$ to demonstrate the improvement in the adjusted fourth moment estimator β_1^{mom} over the Copas estimator which has a jump discontinuity. Simulations for other slope estimators have been reported by Hussin (2004). We used an exponential distribution for X with $\mu_X = 10$, and set $\beta_1 = 1$ and $\beta_1 = 0$. The values for the error standard deviations were $(\sigma_\delta,\ \sigma_\tau) \in \{1,2,3,4\}\times\{1,2,3,4\}$, the sample size was $n = 100$ and the number of replications $R = 1000$. We report in Tables 6, 7, and 8 the MSE and the Bias for the estimators $\{\beta_1^{ver}, \beta_1^{hor}, \beta_1^{per}, \beta_1^{gm}, \beta_1^{mom}, \beta_1^{md}\}$ for $(\sigma_\delta,\ \sigma_\tau) \in \{(1,2),(1,3),(1,4)\}$. Similar results hold for $(\sigma_\delta,\ \sigma_\tau) \in \{(2,1),(3,1),(4,1)\}$ Note that in each case the adjusted fourth moment estimator β_1^{mom} is more efficient than the Copas estimator. To see this we compare the pairs of values (MSE, Bias) in the three tables. For β_1^{mom} these are $\{(1.001,-0.830),(2.786,-1.807),(5.717,-2.813)\}$ and for Copas these are $\{(2.378,-2.410),(8.769,-7.347),(23.018,-13.848)\}$. In practice, the researcher may not know which of $\{\sigma_\tau^2, \sigma_\delta^2\}$ is larger. If he does, then he may choose either of $\{\beta_1^{ver}, \beta_1^{hor}\}$ with β_1^{hor} favored when σ_δ^2 is much bigger than σ_τ^2. A fairer comparison is to use OLS$(y|x)$ and OLS$(x|y)$ each 50% of the time. Thus in the Tables we report the average for the MSE and the average of the absolute deviation of the biases for the two OLS estimators. These average (MSE, Bias) values from the tables are $\{(1.336,2.518),(4.847,4.831),(12.46,7.858)\}$ showing the improved efficiency of β_1^{mom}. As anticipated the minimum deviation estimator β_1^{ml} achieves further improvement in reduction of (MSE, Bias) with values $\{(0.646,-1.336),(2.309,-3.584),(5.578,-6.288)\}$.

Table 6

X is $Exp(10), \beta_1 = 1, \beta_0 = 0, R = 1000, n = 100$

$(\sigma_\tau = 1, \sigma_\delta = 2)$

OLS^* reports average MSE and average absolute Bias for $\{\beta_1^{ver}, \beta_1^{hor}\}$

	$MSE * 10^{-3}$	$\%Bias$	λ	θ_λ
β_1^{ver}	2.001	−3.843	1.000	46.12
OLS^*	1.336	2.518	NA	NA
β_1^{hor}	0.670	1.193	0.000	136.12
β_1^{per}	0.688	−1.396	0.507	89.99
β_1^{gm}	0.653	−1.360	0.500	90.78
β_1^{mom}	1.001	−0.830	0.339	108.27
β_1^{cop}	2.378	−2.410	0.651	74.47
β_1^{md}	0.646	−1.336	0.497	91.06

Table 7

X is $Exp(10), \beta_1 = 1, \beta_0 = 0, R = 1000, n = 100$

$(\sigma_\tau = 1, \sigma_\delta = 3)$

OLS^* reports average MSE and average absolute Bias for $\{\beta_1^{ver}, \beta_1^{hor}\}$

	$MSE * 10^{-3}$	$\%Bias$	λ	θ_λ
β_1^{ver}	8.370	−8.459	1.000	47.53
OLS^*	4.847	4.831	NA	NA
β_1^{hor}	1.324	1.203	0.000	137.53
β_1^{per}	2.688	−3.954	0.520	89.60
β_1^{gm}	2.423	−3.760	0.500	92.19
β_1^{mom}	2.786	−1.807	0.318	110.94
β_1^{cop}	8.769	−7.347	0.848	58.14
β_1^{md}	2.309	−3.584	0.490	93.196

Table 8
X is $Exp(10), \beta_1 = 1, \beta_0 = 0, R = 1000, n = 100$
$(\sigma_\tau = 1, \sigma_\delta = 4)$
OLS^* reports average MSE and average absolute Bias for $\{\beta_1^{ver}, \beta_1^{hor}\}$

	$MSE * 10^{-3}$	$\%Bias$	λ	θ_λ
β_1^{ver}	22.791	−14.376	1.000	49.43
OLS^*	12.46	7.858	NA	NA
β_1^{hor}	2.134	1.339	0.000	139.43
β_1^{per}	7.406	−7.480	0.539	89.95
β_1^{gm}	6.242	−6.880	0.500	94.08
β_1^{mom}	5.717	−2.813	0.286	114.51
β_1^{cop}	23.018	−13.848	0.950	52.71
β_1^{md}	5.578	−6.288	0.480	96.04

9 Summary

We have modified the fourth moment estimator of the slope from Gillard and Iles (2005) to show how to remove the jump discontinuity in the estimator given by Copas (1972). We show how the moment estimators $\{\beta_1^{mom}, \widetilde{\sigma}_\delta^2, \widetilde{\sigma}_\tau^2\}$ can be used to determine an MLE estimator which surprisingly is the original moment estimator of the slope. Our simulations support our claim that both $\{\beta_1^{mom}, \beta_1^{md}\}$ are more efficient than the average of the OLS estimators.

10 References

Adcock, R. J. (1878). A problem in least-squares. The Analyst 5:53-54.

Carroll, R. J., Ruppert, D., Stefanski, L. A., Crainiceanu, C. M. (2006). *Measurement Error in Nonlinear Models - A Modern Perspective*, *Second Edition.* Boca Raton: Chapman & Hall/CRC.

Copas, J. (1972). The likelihood surface in the linear functional relationship problem. Journal of the Royal Statistical Society. Series B (Methodological) 34:274-278.

Deming, W. E. (1943). *Statistical Adjustment of Data.* New York: Wiley.

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

Gillard J.,Iles T. (2005). Method of moments estimation in linear regression with errors in both variables. Cardiff University School of Mathematics Technical Report.

Gillard, J., Iles T. (2009). Methods of fitting straight lines where both variables are subject to measurement error. Current Clinical Pharmacology 4:164-171.

Hussin, A.G. (2004). Numerical comparisons for various estimators of slope parameters for unreplicated linear functional model.Matematika 20:19-30.

Lindley, D., El-Sayyad, M. (1968). The Bayesian estimation of a linear functional relationship. Journal of the Royal Statistical Society Series B (Methodological) 30:190-202

Linnet, Kristian (1993). Evaluation of regression procedures for methods comparison studies. Clinical Chemistry 39:432-432.

Linnet, Kristian (1999). Necessary sample size for method comparison studies based on regression analysis Clinical Chemistry 45:882-894.

Madansky, A. (1959). The fitting of straight Lines when both variables are subject to error. Journal of American Statistical Association 54:173-205.

O'Driscoll, D., Ramirez, D., Schmitz, R. (2008). Minimizing oblique errors for robust estimation. Irish. Math. Soc. Bulletin 62*:71-78.*

Riggs, D., Guarnieri, J., Addelman, S. (1978). Fitting straight lines when both variables are subject to error. Life Sciences 22:1305-1360.

Sprent, P. (1970) The saddle point of the likelihood surface for a linear functional relationship. Journal of the Royal Statistical Society. Series B (Methodological) 32:432-434.

Solari, M. (1969) The "Maximum Likelihood Solution" of the problem of estimating a linear functional relationship. Journal of the Royal Statistical Society. Series B (Methodological) 31:372-375.