Investigation of Non-normality in a Simple Errors-in-variables Model

STAT3799 Directed studies in statistics

Lee Chun Yin
3035469140

Supervisor: Dr. Raymond W.L. Wong

April 2021

Abstract

In a classical linear regression setting, we often assume that the explanatory variable is nonrandom without any measurement error, and that the errors are normally distributed. However, this may not be the case in real-life applications, where measurement errors may exist, and the errors may be heavy-tailed or skewed. We use the computer simulation technique to demonstrate the impacts of non-normality in the *errors-in-variables* model. We present numerical results from simulations based on normal, Student's t and χ^2 distributions on the ordinary least squares and method of moments estimation of regression slope parameter β and residual variance σ_ϵ^2.

1 Introduction

Consider the problem of regression through the origin with only one explanatory variable:

$$y = \beta x + \epsilon \tag{1}$$

In real life applications, usually we will first obtain pairs of observations $(\tilde{x}_i, \tilde{y}_i)$, then apply a model by performing linear regression on the data. However, we only have observations on the observed $\tilde{x}_i = x_i + u_i$ and $\tilde{y}_i = y_i + v_i$, which have some additive error compared to the true x and y that these model assumptions are based on. We further assume that the measurement errors u_i and v_i have mean zero and constant variances, and that the measurement error is uncorrelated and independent to the true values x and y. This gives rise to the *errors-in-variables model.* This imposes a different problem from the classical linear regression model, because the classical model assumes that the observed $\tilde{x}$ is nonrandom, that we have access to the true value of explanatory variable x without any error.

Furthermore, in the classical linear regression model, we often assume that the observed dependent variable y is subject to some additive residual ϵ following $N(0, \sigma^2)$. However, the normality assumption often does not hold for real life datasets. For example, when the errors in the dataset have heavy-tails and/or have skewed shapes, then the normal assumption may not be appropriate. For instance, when dealing with datasets with heavy-tailed errors, one of the practices is to assume t-distributed errors instead of normal-distributed errors, as the light-tailedness of the normal distribution essentially implies that we assume that large errors occur with very low probability, which may not be true in datasets of poorer quality. To make the situation more complicated, in practice non-normality in errors may happen in both the residual ϵ and the measurement error u or v. Thus, there is a need to investigate the impacts of non-normality on the *errors-in-variables model.*

In this work, we investigate how non-normality in both the measurement error of explanatory variable x and the residual of the dependent variable y affects the estimation of the regression coefficient β and the estimation of variance of error σ_ϵ^2. We first perform a literature review on existing results on the *errors-in-variables model.* Then we will describe in detail the methodology, which is the computer simulation technique used to produce results. Finally, we will present the results and findings from the computer simulation experiments. The code used in this work can be found in the appendix.

2 Literature review

We mainly refer to the lecture notes written by Pischke [1] for the errors-in-variables model.

Suppose we wish to estimate the relationship $y = \beta x + \epsilon$, but we only have data on $\tilde{x} = x + u$. Also, let's further assume that $\sigma_v^2 = 0$, i.e. there is only measurement error in x.

If we substitute $\tilde{x} = x + u$ into $y = \beta x + \epsilon$, we obtain:

$$y = \beta(\tilde{x} - u) + \epsilon = \beta\tilde{x} + (\epsilon - \beta u) \tag{2}$$

As the measurement error in x becomes part of the residual error term in the model, the exogeneity assumption of the Gauss-Markov theorem is violated as $cov(u, \tilde{x}) \neq 0$. Thus, the ordinary least-squares (OLS) estimator of β may not be the *best linear unbiased estimator* (BLUE). Unlike the case under the presence of measurement error, the OLS and MME estimators are different. In fact, We will see that the OLS estimators of β and σ_ϵ^2 are biased. In order to obtain unbiased and consistent estimates, we would have to resort to the method of moments (MME) estimators instead. Furthermore, although the OLS estimator is the same as the maximum likelihood estimator under normal-distributed errors, this may not be the case under non-normal errors.

2.1 OLS and MME for β

Suppose we use the ordinary least-squares (OLS) estimator for β:

$$\hat{\beta}_{OLS} = \frac{cov(\tilde{x}, y)}{var(\tilde{x})} = \frac{cov(x + u, \beta x + \epsilon)}{var(x + u)} \tag{3}$$

Because ϵ, u and x are independent to each other, we can obtain the limit of $\hat{\beta}_{OLS}$

$$\text{plim } \hat{\beta}_{OLS} = \frac{\beta\sigma_x^2}{\sigma_x^2 + \sigma_u^2} = \lambda\beta \tag{4}$$

where

$$\lambda \equiv \frac{\sigma_x^2}{\sigma_x^2 + \sigma_u^2} \tag{5}$$

This λ is also called the reliability ratio or signal-to-variance ratio.

Therefore, we can see that the OLS estimator $\hat{\beta}$ is biased towards zero because $0 < \lambda < 1$. The sign of the bias depends on the sign of the true β.

In order to obtain consistent estimates for β, we can use the MME estimator instead. Suppose we have some prior knowledge on the measurement errors and have obtained the value of σ_x^2, σ_u^2 or λ. Then we can apply the appropriate adjustment for the bias in $\hat{\beta}_{OLS}$ as $\sigma^2_{\tilde{x}} = \sigma^2_x + \sigma^2_u$, and we can estimate $\sigma^2_{\tilde{x}}$ using $var(\tilde{x})$, which can be measured directly from the observation data.

Some MME estimators for β using the first and second moments alone are shown below [2]:

Table 1: Various MME estimators for β

Assumption	Method of Moments Estimator
σ^2_x known	$\frac{cov(\tilde{x},y)}{\sigma^2_x}$
σ^2_u known	$\frac{cov(\tilde{x},y)}{var(\tilde{x})-\sigma^2_u}$
Reliability ratio λ known	$\frac{cov(\tilde{x},y)}{\lambda var(\tilde{x})}$

There are practical methods available to obtain information on the measurement errors. For example, to obtain the value of σ^2_u, we can perform repeated measurements [2]. To obtain the reliability ratio λ, we can use methods including "intraclass correlation via an internal replication study" [2].

2.2 OLS and MME for σ^2_ϵ

The usual way in OLS to estimate σ^2_ϵ is to calculate $\hat{\sigma^2_\epsilon}$ as the sum of squares of the residuals divided by the degrees of freedom $n-1$. To find out what happens to $\hat{\sigma^2_\epsilon}$, we can first look at what happens to the estimated residual first [1]:

$$\begin{aligned}\hat{\epsilon} &= y - \hat{\beta}_{OLS}\tilde{x} \\ &= y - \hat{\beta}_{OLS}\,(x+u) \\ &= \epsilon - (y - \beta x) + y - \hat{\beta}_{OLS}\,x - \hat{\beta}_{OLS}\,u \\ &= \epsilon + (\beta - \hat{\beta}_{OLS})\,x - \hat{\beta}_{OLS}\,u\end{aligned} \tag{6}$$

Thus, we can see that the true error term is obfuscated by two additional sources of error.

We can further obtain the limit of the OLS variance estimator, as we assumed earlier that ϵ, x and u are uncorrelated:

$$\text{plim } \hat{\sigma^2_\epsilon}_{OLS} = \sigma^2_\epsilon + (1-\lambda)^2\beta^2\sigma^2_x + \lambda^2\beta^2\sigma^2_u \tag{7}$$

In order to obtain the MME estimator for σ^2_ϵ, similar to the case for $\hat{\beta}_{MME}$, if we have some prior knowledge on the measurement error, we can obtain consistent estimates for σ_x^2, σ_u, λ and β. We can obtain a consistent estimate for β using the MME. By substituting $\hat{\beta}_{MME}$ into the expression for $\hat{\sigma^2_\epsilon}_{OLS}$ [2], after rearranging we have:

$$\hat{\sigma^2_\epsilon}_{MME} = \hat{\sigma^2_\epsilon}_{OLS} - (1-\lambda)^2(\hat{\beta}_{MME})^2\sigma_x^2 - \lambda^2(\hat{\beta}_{MME})^2\sigma_u^2 \tag{8}$$

Similarly, based on the relation $\sigma_{\tilde{x}}^2 = \sigma_x^2 + \sigma_u^2$, if σ_u^2 is known, we can obtain:

$$\hat{\sigma^2_\epsilon}_{MME} = \hat{\sigma^2_\epsilon}_{OLS} - (1-\lambda)^2(\hat{\beta}_{MME})^2(var(\tilde{x}) - \sigma_u^2) - \lambda^2(\hat{\beta}_{MME})^2\sigma_u^2 \tag{9}$$

3 Methodology

3.1 Procedures

The computer simulation method was employed in order to investigate the impacts of non-normal errors on the OLS and MME estimates.

For all simulations, we use a common β_{truth}, n, x_{lo}, x_{hi}. β_{truth} is the underlying ground truth β of the model $y = \beta x + \epsilon$. n is the number of observations used in each simulation trial. x_{lo} and x_{hi} are the lower bound and upper bound of the x being sampled respectively.

In order to produce comparable results when using different error distributions, for each experiment, we first fix a certain σ_u^2 and σ_ϵ^2. Afterwards, for the measurement error u, we choose a distribution where the measurement errors u_i are sampled from. The parameters of this distribution are chosen such that it has variance equal to σ_u^2. Similarly, for the residual error ϵ, we choose a distribution where the residual errors ϵ_i are sampled from. The parameters of this distribution are chosen such that it has variance equal to σ_ϵ^2.

After determining which distributions to use, we perform the observation data generation step. First, we pick n instances of explanatory variable x_i uniformly from the interval $[x_{lo}, x_{hi}]$. Afterwards, for each of the instances x_i, we compute $y_i = \beta_{truth} \cdot x_i + \epsilon_i$, where each ϵ_i is drawn independently from the distribution for residual errors determined for this simulation trial. For each x_i, we also compute the observed explanatory variable $\tilde{x}_i = x_i + u_i$, where measurement error u_i is added for each x_i. Each u_i is drawn independently from the distribution for measurement errors determined for this simulation trial.

After generating the observation data, we perform estimations based on the observation data. We perform both the OLS and MME estimations for β and

σ_ϵ^2. For the MME estimations, we use the case where prior information on the measurement error σ_u^2 is known. The relevant expressions for the estimators are listed below:

$$
\begin{aligned}
\hat{\beta}_{OLS} &= \frac{cov(\tilde{x, y})}{var(\tilde{x})} \\
\hat{\beta}_{MME} &= \frac{var(\tilde{x})}{var(\tilde{x}) - \sigma_u^2} \\
\hat{\sigma_\epsilon^2}_{OLS} &= \frac{\sum \hat{\epsilon}}{n-1} \\
\hat{\sigma_\epsilon^2}_{MME} &= \hat{\sigma_\epsilon^2}_{OLS} - (1-\lambda)^2(\hat{\beta}_{MME})^2(var(\tilde{x}) - \sigma_u^2) - \lambda^2(\hat{\beta}_{MME})^2\sigma_u^2
\end{aligned}
\tag{10}
$$

After computing the estimators, we compute the bias and squared error of each of these estimators when compared to the ground truth β_{truth} and σ_ϵ^2.

For the same measurement error distribution and residual error distribution, we perform the whole observation data generation and estimation procedures for multiple iterations. This is to obtain more reliable conclusions on the mean bias error (MBE) and mean squared error (MSE) of the estimators.

The computer simulation procedures are summarized in Procedure 1.

3.2 Choice of distribution

In order to investigate the impacts on non-normality on the estimators, there is a need to use non-normal distributions to sample the measurement error u and residual error ϵ. In order to satisfy the assumptions of the errors-in-variables model, the distributions used to generate the error terms should have mean zero. Furthermore, in our experiment procedures, we also have to derive the model parameters from a fixed σ_u^2 and σ_ϵ^2. Thus, in order to ease calculation, the distribution chosen should have finite variance and the variance is in a closed form such that the derivation of model parameters from a fixed variance is easy to compute.

Taking the above factors into consideration, we decided to use the following distributions: The normal/Gaussian distribution, the Student's t distribution, and the χ^2 distribution re-centered at mean 0. They are chosen because they demonstrate features not seen in the normal distribution - the Student's t distribution has heavier tails than the normal distribution, and the χ^2 distribution is skewed. They pose substantial differences from the normal distribution, which has light tails and is symmetric at 0.

Procedure 1: Computer simulation procedure

```
Initialize β_truth, n, x_lo, x_hi;
forall σ_u^2 ∈ {list of observation variance} do
    forall σ_ε^2 ∈ {list of residual variance} do
        Initialize MSE_{β̂_OLS} = 0, MSE_{β̂_MME} = 0, MBE_{β̂_OLS} = 0,
         MBE_{β̂_MME} = 0.;
        Initialize MSE_{σ̂²_ε OLS} = 0, MSE_{σ̂²_ε MME} = 0, MBE_{σ̂²_ε OLS} = 0,
         MBE_{σ̂²_ε MME} = 0;
        for 10000 iterations do
            Pick n instances of explanatory variable x_i uniformly from
             the interval [x_lo, x_hi];
            Compute n observations of y_i = β_truth · x_i + ε_i, where ε_i is
             drawn independently from a distribution with variance σ_ε^2;
            Compute n instances of observed explanatory variable
             x̃_i = x_i + u_i, where observation u_i is drawn independently
             from a distribution with variance σ_u^2;
            Calculate β̂_OLS, β̂_MME, σ̂²_ε OLS, σ̂²_ε MME from the
             observations x̃, y, and prior information on measurement
             error σ_u^2;
            Update MSE_{β̂_OLS}, MSE_{β̂_MME}, MBE_{β̂_OLS}, MBE_{β̂_MME},
             MSE_{σ̂²_ε OLS}, MSE_{σ̂²_ε MME}, MBE_{σ̂²_ε OLS}, MBE_{σ̂²_ε MME};
        end
    end
end
```

We will then perform the simulations using these distributions as the underlying sampling distributions of the measurement error u and residual error ϵ, as described in the previous subsection. For simplicity, we also chose the same type of distribution for both the measurement error u and residual error ϵ.

3.3 Experiment parameters used

To produce the results in the subsequent section, we used the following parameters: $n = 30$, $\beta_{truth} = 2$, $x_{lo} = -20$, $x_{hi} = 20$. For the simulation procedures, we used $\sigma_u^2 \in \{0, 1.5, 2, 2.5, 3.0, 4.5, 6.0, 8.0, 10.0\}$, and $\sigma_\epsilon^2 \in \{1.5, 2, 2.5\}$. We are unable to experiment with variances from $(0, 1]$ because the variance of the Student's t distribution is only defined when $\sigma^2 > 1$.

For each combination of measurement error distribution and residual error distribution, we performed $10,000$ iterations to produce the MBE and MSE results.

3.4 Implementation details

The codes were implemented in Python in a Jupyter notebook. We used the sampling distribution implementations provided in the `numpy.random` package. The detailed implementation and source code can be found in the appendix.

4 Results and discussions

We split this section into several subsections.

First, we describe the results for the estimation of β. We consider both the MBE and MSE for the OLS and MME estimators of β. We compare the simulation outcomes for different distributions of residual error for when there is absence of measurement error (i.e. $\sigma_u^2 = 0$). For the cases under the presence of measurement error (i.e. $\sigma_u^2 > 0$), we compare the MBE and MSE for different distributions (Normal, Student's t, χ^2) while maintaining the same variance.

Afterwards, we similarly describe the results for the estimation of σ_ϵ^2. We consider both the MBE and MSE for the OLS and MME estimators of σ^2. We similarly compare the simulation outcomes under the cases of absence of measurement error, and cases under the presence of measurement error of different distributions respectively.

4.1 Estimation of β

We use the following estimators of β and apply them to the generated observation data $(\tilde{x}_i, y_i)$:

$$\begin{aligned} \hat{\beta}_{OLS} &= \frac{cov(\tilde{x, y})}{var(\tilde{x})} \\ \hat{\beta}_{MME} &= \frac{var(\tilde{x})}{var(\tilde{x}) - \sigma_u^2} \end{aligned} \tag{11}$$

4.1.1 Absence of measurement error

Under the absence of measurement error, the OLS and MME estimators of β are the same.

Table 2 summarizes the MBE of β when $\sigma_u^2 = 0$ for different error distri-

Table 2: Estimation MBE of β when $\sigma_u^2 = 0$, under different distributions and variance of ϵ.

	Normal	Student's t	χ^2
$\sigma_\epsilon^2 = 1.5$	-0.000188	-0.000124	0.000201
$\sigma_\epsilon^2 = 2.0$	-0.000114	0.000175	0.000259
$\sigma_\epsilon^2 = 2.5$	-0.000217	-0.000189	0.000073

butions for ϵ. There is no clear trend of increasing/decreasing pattern of the MBE for each of the distributions. This is mostly because the OLS estimator is an unbiased estimator of β, making the bias to be centered at zero, leading to the signs of bias to be different between simulation trials, which cancel out each other in the calculation of MBE.

Table 3: Estimation MSE of β when $\sigma_u^2 = 0$, under different distributions and variance of ϵ.

	Normal	Student's t	χ^2
$\sigma_\epsilon^2 = 1.5$	0.000346	0.000355	0.000358
$\sigma_\epsilon^2 = 2.0$	0.000477	0.000459	0.000464
$\sigma_\epsilon^2 = 2.5$	0.000583	0.000589	0.000588

Table 3 summarizes the MSE of β when $\sigma_u^2 = 0$ for different error distributions. There is a clear trend of increasing of the MSE for each of the distributions as σ_ϵ^2 increases, which is intuitive as the observation data becomes dirtier as σ_ϵ^2 increases. We can also observe that the MSE for Student's t-distributed errors and χ^2-distributed errors are greater in most cases.

4.1.2 Normal-distributed measurement error and normal-distributed residual error

Under the presence of measurement error, the OLS and MME estimators of β are different. We would expect the MME estimator to have lower absolute bias and lower MSE because under the presence of measurement error, the MME estimator is consistent but the OLS estimator is inconsistent.

Table 4 and Table 5 summarize the simulated MBE and MSE of the estimation of β respectively, under normal-distributed errors.

We first focus on the observations of the MBE and MSE of $\hat{\beta}_{OLS}$. Firstly, we observe that the MBE of $\hat{\beta}_{OLS}$ is negative. This is result is expected because plim $\hat{\beta}_{OLS} = \frac{\beta\sigma_x^2}{\sigma_x^2+\sigma_u^2} = \lambda \cdot \beta_{truth}$, making $\hat{\beta}_{OLS}$ to be biased towards zero. In this simulation, $\beta_{truth} = 2 > 0$. Thus, the bias of $\hat{\beta}_{OLS}$ is expected to be

Table 4: Estimation MBE of $\hat{\beta}_{OLS}$ and $\hat{\beta}_{MME}$ under normal-distributed measurement error with different variances σ_u^2 and residual error with different variances σ_ϵ^2.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$
$\sigma_u^2 = 1.5$	-0.018406	0.001800	-0.019667	0.000505	-0.019292	0.000895
$\sigma_u^2 = 2.0$	-0.025018	0.001851	-0.024511	0.002371	-0.025436	0.001416
$\sigma_u^2 = 2.5$	-0.031520	0.001958	-0.029421	0.004139	-0.031215	0.002285
$\sigma_u^2 = 3.0$	-0.036707	0.003436	-0.036961	0.003156	-0.037188	0.002923
$\sigma_u^2 = 4.5$	-0.054875	0.004901	-0.056101	0.003559	-0.055655	0.004057
$\sigma_u^2 = 6.0$	-0.073502	0.005575	-0.074400	0.004531	-0.072289	0.006886
$\sigma_u^2 = 8.0$	-0.097213	0.007125	-0.097049	0.007313	-0.096943	0.007427
$\sigma_u^2 = 10.0$	-0.120748	0.008388	-0.119521	0.009902	-0.119075	0.010326

Table 5: Estimation MSE of $\hat{\beta}_{OLS}$ and $\hat{\beta}_{MME}$ under normal-distributed measurement error with different variances σ_u^2 and residual error with different variances σ_ϵ^2.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$
$\sigma_u^2 = 1.5$	0.002026	0.001781	0.002168	0.001873	0.002304	0.002029
$\sigma_u^2 = 2.0$	0.002732	0.002267	0.002830	0.002396	0.002987	0.002506
$\sigma_u^2 = 2.5$	0.003544	0.002801	0.003559	0.002964	0.003688	0.002967
$\sigma_u^2 = 3.0$	0.004319	0.003334	0.004455	0.003456	0.004653	0.003648
$\sigma_u^2 = 4.5$	0.007088	0.004862	0.007457	0.005112	0.007360	0.005045
$\sigma_u^2 = 6.0$	0.010532	0.006487	0.010886	0.006739	0.010648	0.006833
$\sigma_u^2 = 8.0$	0.016103	0.009102	0.016091	0.009100	0.016123	0.009138
$\sigma_u^2 = 10.0$	0.022169	0.011236	0.022276	0.011799	0.022216	0.011831

negative, which is consistent with the simulation findings. Second, we observe that the MBE of $\hat{\beta}_{OLS}$ decreases as σ_u^2 increases. This is also consistent with the literature as $\lambda \equiv \frac{\sigma_x^2}{\sigma_x^2+\sigma_u^2}$. When σ_u^2 increases, λ decreases, and the bias $(\lambda - 1) \cdot \beta_{truth}$ decreases. Third, we observe that the MSE of $\hat{\beta}_{OLS}$ increases as σ_u^2 increases. Intuitively, this means that the error of the estimation of β increases as the observation data is dirtier.

Now, we turn our attention to the observations of the MBE and MSE of $\hat{\beta}_{MME}$, and compare them with $\hat{\beta}_{OLS}$. Firstly, we observe that the absolute value of MBE and MSE of $\hat{\beta}_{MME}$ are both lower than that of $\hat{\beta}_{OLS}$ at the same σ_u^2 and σ_ϵ^2. This is expected because $\hat{\beta}_{MME}$ is a consistent estimator of β while $\hat{\beta}_{OLS}$ is not. Second, we observe that the MBE of $\hat{\beta}_{MME}$ is positive and in general increases as σ_u^2 increases. Third, we observe that the MSE of $\hat{\beta}_{MME}$

increases as σ_u^2 increases. This is similar to the case of $\hat{\beta}_{OLS}$.

4.1.3 Student's t-distributed measurement error and Student's t-distributed residual error

We perform the similar simulations for Student's t distributed measurement error and residual error for different variances. Elementary statistical theory tells us that the variance for a Student's t distribution with $\nu > 2$ degrees of freedom has variance equal to $\frac{\nu}{\nu-2}$. Taking inverse of the equation $\sigma^2 = \frac{\nu}{\nu-2}$ gives $\nu = 2 \cdot \frac{\sigma^2}{\sigma^2-1}$. Note that the range of σ_u^2 and σ_ϵ^2 are specially chosen such that $\nu > 2$ because the Student's t distribution only has finite variance when $\nu > 2$. Thus, in the tables shown below, when we write $\sigma^2 = \delta$, it means the simulation trial was performed using the Student's t distribution with $2 \cdot \frac{\delta}{\delta-1}$ degrees of freedom.

Table 6 and Table 7 summarize the simulated MBE and MSE of the estimation of β respectively, under Student's t-distributed errors.

Table 6: Estimation MBE of $\hat{\beta}_{OLS}$ and $\hat{\beta}_{MME}$ under Student's t-distributed measurement error with different variances σ_u^2 and residual error with different variances σ_ϵ^2.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$
$\sigma_u^2 = 1.5$	-0.018811	0.001391	-0.019199	0.000988	-0.018163	0.002050
$\sigma_u^2 = 2.0$	-0.024563	0.002314	-0.024635	0.002238	-0.024720	0.002157
$\sigma_u^2 = 2.5$	-0.031062	0.002428	-0.031050	0.002448	-0.030458	0.003071
$\sigma_u^2 = 3.0$	-0.034826	0.005397	-0.035187	0.005009	-0.035580	0.004603
$\sigma_u^2 = 4.5$	-0.045805	0.014550	-0.048785	0.011340	-0.045181	0.015206
$\sigma_u^2 = 6.0$	-0.054096	0.026632	-0.054759	0.025926	-0.054605	0.026035
$\sigma_u^2 = 8.0$	-0.061836	0.046525	-0.060023	0.048540	-0.061507	0.046848
$\sigma_u^2 = 10.0$	-0.066070	0.070679	-0.067018	0.069615	-0.065713	0.071101

We first focus on the observations based solely on the Student's t-distributed error simulations. This will be followed by comparing the results in normal-distributed errors and Student's t-distributed errors.

Firstly, similar to the normal-distributed errors case, we observe that the MBE of $\hat{\beta}_{OLS}$ is negative and decreasing in σ_u^2, and that the MBE of $\hat{\beta}_{MME}$ is positive and increasing in σ_u^2. The MSE of $\hat{\beta}_{OLS}$ and $\hat{\beta}_{MME}$ are both increasing in σ_u^2.

When comparing $\hat{\beta}_{OLS}$ and $\hat{\beta}_{MME}$, we find that switching from $\hat{\beta}_{OLS}$ to $\hat{\beta}_{MME}$ still provides a decrease in absolute MBE or MSE, but the decrease is not

Table 7: Estimation MSE of $\hat{\beta}_{OLS}$ and $\hat{\beta}_{MME}$ under Student's t-distributed measurement error with different variances σ_u^2 and residual error with different variances σ_ϵ^2.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$
$\sigma_u^2 = 1.5$	0.002112	0.001854	0.002179	0.001905	0.002279	0.002051
$\sigma_u^2 = 2.0$	0.002912	0.002481	0.003040	0.002609	0.003132	0.002699
$\sigma_u^2 = 2.5$	0.004094	0.003409	0.004211	0.003528	0.004270	0.003635
$\sigma_u^2 = 3.0$	0.005661	0.004913	0.005470	0.004671	0.005644	0.004827
$\sigma_u^2 = 4.5$	0.009523	0.008639	0.010456	0.009277	0.009140	0.008293
$\sigma_u^2 = 6.0$	0.014130	0.013794	0.013723	0.013223	0.013692	0.013231
$\sigma_u^2 = 8.0$	0.017969	0.019371	0.016112	0.017761	0.017037	0.018482
$\sigma_u^2 = 10.0$	0.019388	0.024422	0.020242	0.025071	0.018737	0.023675

that large when compared to the normal-distributed errors case. This situation is much more apparent in large values of σ_u^2. For MBE, when $\sigma_u^2 = 10.0$, we observe that the absolute value of the MBE of $\hat{\beta}_{MME}$ is in fact greater than that of $\hat{\beta}_{OLS}$. Similarly for MSE, when $\sigma_u^2 \in \{8.0, 10.0\}$, we observe that the MSE of $\hat{\beta}_{MME}$ is greater than that of $\hat{\beta}_{OLS}$.

To compare the findings between the normal-distributed errors case and Student's t-distributed errors case, it is useful to take the cell difference between the estimation MBE in Tables 4 and 6, and estimation MSE in Tables 5 and 7. We define the difference here as (value from Student's t) $-$ (value from normal). These differences are tabulated in Tables 8 and 9 respectively.

Table 8: Difference between Student's t error case and normal-distributed error case: Estimation MBE of $\hat{\beta}_{OLS}$ and $\hat{\beta}_{MME}$ under different σ_u^2 and σ_ϵ^2.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$
$\sigma_u^2 = 1.5$	-0.000405	-0.000409	0.000469	0.000483	0.001129	0.001155
$\sigma_u^2 = 2.0$	0.000455	0.000463	-0.000124	-0.000133	0.000717	0.000742
$\sigma_u^2 = 2.5$	0.000458	0.000470	-0.001629	-0.001691	0.000757	0.000786
$\sigma_u^2 = 3.0$	0.001881	0.001961	0.001774	0.001853	0.001608	0.001680
$\sigma_u^2 = 4.5$	0.009069	0.009649	0.007316	0.007781	0.010474	0.011148
$\sigma_u^2 = 6.0$	0.019406	0.021057	0.019641	0.021394	0.017684	0.019149
$\sigma_u^2 = 8.0$	0.035377	0.039400	0.037027	0.041226	0.035436	0.039422
$\sigma_u^2 = 10.0$	0.054678	0.062291	0.052503	0.059713	0.053362	0.060775

It should be noted that when the variance of a Student's t distribution is low, the number of degrees of freedom of the distribution is high. From elementary statistical theory we know that as the number of degrees of freedom of a

Table 9: Difference between Student's t error case and normal-distributed error case: Estimation MSE of $\hat{\beta}_{OLS}$ and $\hat{\beta}_{MME}$ under different σ_u^2 and σ_ϵ^2.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$
$\sigma_u^2 = 1.5$	0.000086	0.000073	0.000011	0.000031	-0.000024	0.000022
$\sigma_u^2 = 2.0$	0.000181	0.000214	0.000210	0.000212	0.000145	0.000193
$\sigma_u^2 = 2.5$	0.000550	0.000608	0.000652	0.000565	0.000582	0.000668
$\sigma_u^2 = 3.0$	0.001342	0.001579	0.001015	0.001215	0.000990	0.001178
$\sigma_u^2 = 4.5$	0.002435	0.003777	0.002999	0.004165	0.001780	0.003248
$\sigma_u^2 = 6.0$	0.003597	0.007307	0.002836	0.006484	0.003044	0.006398
$\sigma_u^2 = 8.0$	0.001866	0.010269	0.000021	0.008661	0.000914	0.009344
$\sigma_u^2 = 10.0$	-0.002781	0.013186	-0.002034	0.013272	-0.003479	0.011843

Student's t distribution increases, it converges to the normal distribution. Thus, for some low fixed variance, we should expect similar results from a Student's t distribution and normal distribution with that variance.

We first focus on the comparisons in MBE. For $\hat{\beta}_{OLS}$, we observe that the difference is first close to zero, then positive and increasing. As the bias of $\hat{\beta}_{OLS}$ is negative in both Student's t-distributed errors and normal-distributed errors, when the difference in MBE is positive, it means that $\hat{\beta}_{OLS}$ under Student's t errors is closer to zero. We have similar findings for $\hat{\beta}_{MME}$, where the difference is also increasing throughout as σ_u^2 is increased. As the MBE for $\hat{\beta}_{MME}$ is positive in both Student's t errors and normal errors, this means that the bias for $\hat{\beta}_{MME}$ under Student's t errors is greater and further away from zero.

We then focus on the comparisons in MSE. For both $\hat{\beta}_{OLS}$ and $\hat{\beta}_{MME}$, we observe that the difference is mostly positive, meaning that the MSE under Student's t-distributed errors is in general greater than that under normal-distributed errors. For $\hat{\beta}_{OLS}$, this difference first increases to a certain point (around at $\sigma_u^2 = 6.0$), then decreases. For $\hat{\beta}_{MME}$, this differences increases throughout as σ_u^2 is increased.

4.1.4 χ^2-distributed measurement error

We perform the similar simulations for χ^2 distributed measurement error and residual error for different variances. From elementary statistical theory, a χ^2 distribution with $k > 0$ degrees of freedom has mean k and variance $2k$. To let the distribution follow the model assumptions to have mean zero, we subtract k from each value drawn from the distribution. Thus, when we write $\sigma^2 = \delta$, it means the simulation trial was performed using the χ^2 distribution with $\frac{\delta}{2}$ degrees of freedom minus $\frac{\delta}{2}$ to recenter the mean to zero.

Table 10 and Table 11 summarize the simulated MBE and MSE of the estimation of β respectively, under χ^2-distributed errors.

Table 10: Estimation MBE of $\hat{\beta}_{OLS}$ and $\hat{\beta}_{MME}$ under re-centered χ^2-distributed measurement error with different variances σ_u^2 and residual error with different variances σ_ϵ^2.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$
$\sigma_u^2 = 1.5$	-0.018624	0.001580	-0.017994	0.002220	-0.018908	0.001285
$\sigma_u^2 = 2.0$	-0.024808	0.002076	-0.025817	0.001020	-0.024370	0.002517
$\sigma_u^2 = 2.5$	-0.030362	0.003170	-0.031010	0.002494	-0.030079	0.003481
$\sigma_u^2 = 3.0$	-0.036793	0.003351	-0.037685	0.002393	-0.037465	0.002621
$\sigma_u^2 = 4.5$	-0.054270	0.005575	-0.055211	0.004503	-0.056487	0.003124
$\sigma_u^2 = 6.0$	-0.072231	0.006964	-0.072339	0.006804	-0.072609	0.006531
$\sigma_u^2 = 8.0$	-0.097065	0.007307	-0.097076	0.007311	-0.096631	0.007895
$\sigma_u^2 = 10.0$	-0.118317	0.011277	-0.116507	0.013404	-0.118055	0.011505

Table 11: Estimation MSE of $\hat{\beta}_{OLS}$ and $\hat{\beta}_{MME}$ under re-centered χ^2-distributed measurement error with different variances σ_u^2 and residual error with different variances σ_ϵ^2.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$
$\sigma_u^2 = 1.5$	0.002205	0.001958	0.002274	0.002053	0.002460	0.002206
$\sigma_u^2 = 2.0$	0.002944	0.002501	0.003080	0.002584	0.003201	0.002792
$\sigma_u^2 = 2.5$	0.003712	0.003059	0.003838	0.003142	0.003863	0.003232
$\sigma_u^2 = 3.0$	0.004493	0.003508	0.004917	0.003890	0.004644	0.003598
$\sigma_u^2 = 4.5$	0.007488	0.005382	0.007612	0.005388	0.007981	0.005624
$\sigma_u^2 = 6.0$	0.010957	0.007206	0.011047	0.007281	0.011212	0.007411
$\sigma_u^2 = 8.0$	0.016616	0.009711	0.016813	0.009960	0.016772	0.010004
$\sigma_u^2 = 10.0$	0.022696	0.012718	0.022178	0.012613	0.022977	0.013127

For the MBE and MSE, we have similar observations to the normal-distributed errors and Student's t-distributed errors cases, where the MBE of $\hat{\beta}_{OLS}$ is negative and decreasing in σ_u^2, and the MBE of $\hat{\beta}_{MME}$ is positive and increasing in σ_u^2. Similar to before, the MSE of $\hat{\beta}_{OLS}$ and $\hat{\beta}_{MME}$ are both increasing in σ_u^2, and the MSE for $\hat{\beta}_{MME}$ is smaller than that for $\hat{\beta}_{OLS}$.

To compare findings between the normal-distributed errors and χ^2-distributed errors case, we again take the cell difference from the two distributions in their MBE and MSE respectively. The difference here is (value from χ^2) − (value from normal). These differences are tabulated in Tables 12 and 13 respectively.

Table 12: Difference between re-centered χ^2-distributed error case and normal-distributed error case: Estimation MBE of $\hat{\beta}_{OLS}$ and $\hat{\beta}_{MME}$ under different σ_u^2 and σ_ϵ^2.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$
$\sigma_u^2 = 1.5$	-0.000219	-0.000220	0.001673	0.001715	0.000383	0.000390
$\sigma_u^2 = 2.0$	0.000210	0.000225	-0.001305	-0.001351	0.001066	0.001102
$\sigma_u^2 = 2.5$	0.001158	0.001212	-0.001589	-0.001645	0.001136	0.001196
$\sigma_u^2 = 3.0$	-0.000085	-0.000085	-0.000725	-0.000763	-0.000277	-0.000301
$\sigma_u^2 = 4.5$	0.000605	0.000674	0.000890	0.000945	-0.000832	-0.000934
$\sigma_u^2 = 6.0$	0.001271	0.001389	0.002061	0.002273	-0.000320	-0.000354
$\sigma_u^2 = 8.0$	0.000148	0.000181	-0.000027	-0.000002	0.000312	0.000469
$\sigma_u^2 = 10.0$	0.002431	0.002889	0.003014	0.003502	0.001020	0.001180

Table 13: Difference between re-centered χ^2-distributed error case and normal-distributed error case: Estimation MSE of $\hat{\beta}_{OLS}$ and $\hat{\beta}_{MME}$ under different σ_u^2 and σ_ϵ^2.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$
$\sigma_u^2 = 1.5$	0.000180	0.000177	0.000106	0.000180	0.000156	0.000176
$\sigma_u^2 = 2.0$	0.000212	0.000234	0.000250	0.000188	0.000214	0.000286
$\sigma_u^2 = 2.5$	0.000168	0.000259	0.000279	0.000178	0.000175	0.000265
$\sigma_u^2 = 3.0$	0.000174	0.000174	0.000462	0.000434	-0.000009	-0.000050
$\sigma_u^2 = 4.5$	0.000399	0.000521	0.000155	0.000276	0.000621	0.000580
$\sigma_u^2 = 6.0$	0.000425	0.000719	0.000161	0.000542	0.000563	0.000578
$\sigma_u^2 = 8.0$	0.000513	0.000608	0.000723	0.000861	0.000649	0.000866
$\sigma_u^2 = 10.0$	0.000527	0.001482	-0.000098	0.000814	0.000761	0.001296

When comparing findings from χ^2-distributed errors and normal-distributed errors, it should be again noted that the χ^2 distribution re-centered at mean zero converges to a normal distribution with the same variance when the number of degrees of freedom is large.

For the MBE, there seems to be no clear comparison between the normal-distributed and χ^2-distributed cases, because the sign of the difference fluctuates in $\hat{\beta}_{OLS}$, $\hat{\beta}_{MME}$ and different σ_u^2.

For the MSE, the difference is always positive, meaning that the MSE for χ^2-distributed case is always higher than normal-distributed case. The difference seems to be increasing in σ_u^2, but the trend is not that clear.

4.2 Estimation of σ_ϵ^2

We use the following estimators of σ_ϵ^2 and apply them to the generated observation data $(\tilde{x}_i, y_i)$:

$$\begin{aligned}\hat{\sigma_\epsilon^2}_{OLS} &= \frac{\sum \hat{\epsilon}}{n-1} \\ \hat{\sigma_\epsilon^2}_{MME} &= \hat{\sigma_\epsilon^2}_{OLS} - (1-\lambda)^2(\hat{\beta}_{MME})^2(var(\tilde{x}) - \sigma_u^2) - \lambda^2(\hat{\beta}_{MME})^2\sigma_u^2\end{aligned} \tag{12}$$

4.2.1 Absence of measurement error

Under the absence of measurement error, the OLS and MME estimators of σ_ϵ^2 are the same. Tables 14 and 15 summarize respectively the MBE and MSE of σ_ϵ^2 when $\sigma_u^2 = 0$ for different error distributions.

Table 14: Estimation MBE of σ_ϵ^2 when $\sigma_u^2 = 0$, under different distributions and variance of ϵ.

	Normal	Student's t	χ^2
$\sigma_\epsilon^2 = 1.5$	0.002638	-0.005014	0.033716
$\sigma_\epsilon^2 = 2.0$	0.005647	-0.022737	0.018373
$\sigma_\epsilon^2 = 2.5$	-0.014956	-0.005425	0.007073

Table 15: Estimation MSE of σ_ϵ^2 when $\sigma_u^2 = 0$, under different distributions and variance of ϵ.

	Normal	Student's t	χ^2
$\sigma_\epsilon^2 = 1.5$	0.150847	0.344340	1.490189
$\sigma_\epsilon^2 = 2.0$	0.276744	1.504203	1.936359
$\sigma_\epsilon^2 = 2.5$	0.424303	9.540943	2.494662

For MBE, similar to the case of $\hat{\beta}$, there seems to be no clear monotonic trend of the MBE from the simulations performed under different distributions, mostly due to the fact that $\hat{\sigma_\epsilon^2}_{OLS}$ is an unbiased estimator of σ_ϵ^2 under the absence of measurement error. For MSE, we can observe that the MSE for normal-errors is the least out of the distributions in concern, and the MSE for Student's t errors and re-centered χ^2 errors are greater.

4.2.2 Normal-distributed measurement error

We again perform the estimation of σ_ϵ^2 under normal-distributed measurement error and normal-distributed residual error, both using different variances. The simulated MBE and MSE are tabulated in Tables 16 and 17 respectively.

Table 16: Estimation MBE of $\hat{\sigma^2_\epsilon}_{OLS}$ and $\hat{\sigma^2_\epsilon}_{MME}$ under normal-distributed measurement error with different variances σ_u^2 and residual error with different variances σ_ϵ^2.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$
$\sigma_u^2 = 1.5$	5.900404	-0.050233	5.955353	0.012206	5.947017	0.001355
$\sigma_u^2 = 2.0$	7.888963	-0.018928	7.901321	-0.010901	7.879309	-0.025664
$\sigma_u^2 = 2.5$	9.784296	-0.068626	9.802717	-0.071802	9.869316	0.012851
$\sigma_u^2 = 3.0$	11.798199	-0.002881	11.757166	-0.041108	11.793174	-0.002928
$\sigma_u^2 = 4.5$	17.597952	0.047601	17.466602	-0.062097	17.455414	-0.081391
$\sigma_u^2 = 6.0$	23.120196	-0.063780	23.033594	-0.129041	23.132393	-0.083139
$\sigma_u^2 = 8.0$	30.406714	-0.150267	30.219614	-0.342906	30.389711	-0.176612
$\sigma_u^2 = 10.0$	37.647515	-0.097025	37.724051	-0.077960	37.508399	-0.310914

Table 17: Estimation MSE of $\hat{\sigma^2_\epsilon}_{OLS}$ and $\hat{\sigma^2_\epsilon}_{MME}$ under normal-distributed measurement error with different variances σ_u^2 and residual error with different variances σ_ϵ^2.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$
$\sigma_u^2 = 1.5$	38.530776	3.780289	39.882151	4.471836	40.209234	4.935117
$\sigma_u^2 = 2.0$	68.339624	6.241460	69.219747	6.910293	69.534180	7.638777
$\sigma_u^2 = 2.5$	104.624534	9.105856	105.756458	9.905827	108.054738	10.961756
$\sigma_u^2 = 3.0$	151.461698	12.734043	151.520531	13.621373	153.743436	15.163531
$\sigma_u^2 = 4.5$	335.041092	26.578374	331.019140	27.398219	331.821175	28.443274
$\sigma_u^2 = 6.0$	576.830404	45.826542	575.198317	47.750947	581.386469	49.216860
$\sigma_u^2 = 8.0$	995.830908	79.148108	984.510480	79.308313	998.785315	83.400332
$\sigma_u^2 = 10.0$	1520.990888	117.975374	1528.874963	120.145671	1516.974233	125.296896

For $\hat{\sigma^2_\epsilon}_{OLS}$, from the literature review, we know that plim $\hat{\sigma^2_\epsilon}_{OLS} = \sigma_\epsilon^2 + (1-\lambda)^2\beta^2\sigma_x^2 + \lambda^2\beta^2\sigma_u^2$, meaning that there is a positive bias of $\hat{\sigma^2_\epsilon}_{OLS}$. Indeed, we observe that the MBE of $\hat{\sigma^2_\epsilon}_{OLS}$, is positive for all simulated values of σ_u^2 and σ_ϵ^2. The MBE is increasing in σ_u^2 and there is no clear monotonic trend of MBE with σ_ϵ^2, which are both consistent with the expression of additional bias terms in the explicit derivation of plim $\hat{\sigma^2_\epsilon}_{OLS}$.

On the other hand, for $\hat{\sigma^2_\epsilon}_{MME}$, the MBE is negative. The monotonic trend of MBE with σ_u^2 is not that clear. The MSE of $\hat{\sigma^2_\epsilon}_{MME}$ is smaller than that of $\hat{\sigma^2_\epsilon}_{OLS}$, and it is increasing with σ_u^2.

4.2.3 Student's t-distributed measurement error

We perform the similar simulations and estimation of σ_ϵ^2 under the case of Student's t-distributed measurement error and residual error. The MBE and MSE results are tabulated in Tables 18 and 19 respectively:

Table 18: Estimation MBE of $\hat{\sigma^2_\epsilon}_{OLS}$ and $\hat{\sigma^2_\epsilon}_{MME}$ under Student's t-distributed measurement error with different variances σ_u^2 and residual error with different variances σ_ϵ^2.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$
$\sigma_u^2 = 1.5$	5.965444	0.017129	5.949026	0.002994	5.852522	-0.099986
$\sigma_u^2 = 2.0$	7.873037	-0.038917	7.794577	-0.117039	7.837349	-0.073789
$\sigma_u^2 = 2.5$	9.613666	-0.245337	9.697142	-0.162316	9.591657	-0.274076
$\sigma_u^2 = 3.0$	11.240567	-0.587908	11.061106	-0.762236	11.091512	-0.727538
$\sigma_u^2 = 4.5$	14.674276	-3.058067	15.130487	-2.549836	14.349906	-3.392302
$\sigma_u^2 = 6.0$	17.308219	-6.392124	17.300402	-6.380649	17.289769	-6.394398
$\sigma_u^2 = 8.0$	19.391143	-12.403962	18.999681	-12.842478	19.122997	-12.675239
$\sigma_u^2 = 10.0$	20.359688	-19.757988	20.806443	-19.278803	20.337719	-19.788499

Table 19: Estimation MSE of $\hat{\sigma^2_\epsilon}_{OLS}$ and $\hat{\sigma^2_\epsilon}_{MME}$ under Student's t-distributed measurement error with different variances σ_u^2 and residual error with different variances σ_ϵ^2.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$
$\sigma_u^2 = 1.5$	42.851635	7.429251	44.646925	9.417269	47.303721	13.233213
$\sigma_u^2 = 2.0$	90.739181	29.905215	88.083920	28.419856	96.305849	35.838509
$\sigma_u^2 = 2.5$	167.755054	79.503490	181.578188	92.099801	182.441974	94.903265
$\sigma_u^2 = 3.0$	284.186452	168.747879	278.572350	166.286730	276.831587	163.997266
$\sigma_u^2 = 4.5$	621.524421	453.668034	712.260002	535.305676	588.075696	427.984632
$\sigma_u^2 = 6.0$	1059.949532	894.596872	988.210829	815.597278	1118.336892	940.571666
$\sigma_u^2 = 8.0$	1433.607905	1384.842799	1208.084274	1158.996324	1213.814914	1157.835613
$\sigma_u^2 = 10.0$	1430.156559	1641.357813	1618.590469	1813.400195	1429.762567	1631.738065

Similar to the normal case, we find that for OLS the MBE is positive and increasing with σ_u^2, and for MME the MBE is negative and decreasing with

σ_u^2. We find that the improvement in estimation accuracy when using $\hat{\sigma^2_\epsilon}_{MME}$ instead of using $\hat{\sigma^2_\epsilon}_{OLS}$ is not as great as compared to the normal case. Similar to the case of estimating β, we also find that when σ_u^2 increases, the improvement of using $\hat{\sigma^2_\epsilon}_{MME}$ to replace $\hat{\sigma^2_\epsilon}_{OLS}$ decreases. In fact, we see that when $\sigma_u^2 = 10.0$, the MSE of $\hat{\sigma^2_\epsilon}_{MME}$ is greater than that of $\hat{\sigma^2_\epsilon}_{OLS}$.

To compare the findings between normal-distributed errors and Student's t distributed errors, we take differences in a fashion similar to the previous sections. The differences are tabulated in Tables 20 and 21 respectively:

Table 20: Difference between Student's t-distributed error case and normal-distributed error case: Estimation MBE of $\hat{\sigma^2_\epsilon}_{OLS}$ and $\hat{\sigma^2_\epsilon}_{MME}$ under different σ_u^2 and σ_ϵ^2.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$
$\sigma_u^2 = 1.5$	0.065040	0.067362	-0.006327	-0.009212	-0.094495	-0.101341
$\sigma_u^2 = 2.0$	-0.015926	-0.019989	-0.106743	-0.106139	-0.041960	-0.048125
$\sigma_u^2 = 2.5$	-0.170630	-0.176711	-0.105574	-0.090515	-0.277659	-0.286927
$\sigma_u^2 = 3.0$	-0.557633	-0.585027	-0.696060	-0.721128	-0.701662	-0.724610
$\sigma_u^2 = 4.5$	-2.923676	-3.105667	-2.336115	-2.487739	-3.105508	-3.310911
$\sigma_u^2 = 6.0$	-5.811977	-6.328344	-5.733192	-6.251608	-5.842623	-6.311259
$\sigma_u^2 = 8.0$	-11.015571	-12.253694	-11.219933	-12.499572	-11.266714	-12.498626
$\sigma_u^2 = 10.0$	-17.287827	-19.660963	-16.917608	-19.200842	-17.170679	-19.477585

Table 21: Difference between Student's t-distributed error case and normal-distributed error case: Estimation MBE of $\hat{\sigma^2_\epsilon}_{OLS}$ and $\hat{\sigma^2_\epsilon}_{MME}$ under different σ_u^2 and σ_ϵ^2.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$
$\sigma_u^2 = 1.5$	4.320859	3.648962	4.764773	4.945432	7.094487	8.298096
$\sigma_u^2 = 2.0$	22.399557	23.663755	18.864172	21.509563	26.771668	28.199732
$\sigma_u^2 = 2.5$	63.130521	70.397634	75.821730	82.193974	74.387236	83.941509
$\sigma_u^2 = 3.0$	132.724754	156.013836	127.051819	152.665357	123.088151	148.833734
$\sigma_u^2 = 4.5$	286.483328	427.089660	381.240862	507.907457	256.254521	399.541359
$\sigma_u^2 = 6.0$	483.119128	848.770330	413.012512	767.846331	536.950423	891.354806
$\sigma_u^2 = 8.0$	437.776996	1305.694690	223.573794	1079.688011	215.029599	1074.435281
$\sigma_u^2 = 10.0$	-90.834328	1523.382439	89.715506	1693.254524	-87.211667	1506.441169

For MBE, we observe that the differences for both $\hat{\sigma^2_\epsilon}_{OLS}$ and $\hat{\sigma^2_\epsilon}_{MME}$ are both all negative and decreasing in σ_u^2. As the MBE for $\hat{\sigma^2_\epsilon}_{OLS}$ is positive for both normal errors and Student's t errors, this means that the MBE for $\hat{\sigma^2_\epsilon}_{OLS}$ is closer to zero for Student's t errors. However, for $\hat{\sigma^2_\epsilon}_{MME}$, as the MBE is negative for both types of errors, this means that $\hat{\sigma^2_\epsilon}_{MME}$ has a mean bias further away from zero under Student's t errors.

Looking at the MSE, for OLS we observe that the difference is first positive and increasing for smaller σ^2_u, but after $\sigma^2_u = 6.0$ the difference starts to decrease, eventually becoming negative at $\sigma^2_u = 10.0$. For MME we observe that the difference is positive and increasing in σ^2_u throughout. The difference for $\hat{\sigma^2_\epsilon}_{MME}$ is greater than that of $\hat{\sigma^2_\epsilon}_{OLS}$ for the same values of σ^2_u and σ^2_ϵ.

4.2.4 χ^2-distributed measurement error

Lastly, we perform the simulations on estimating σ^2_ϵ under re-centered χ^2 errors. The estimation MBE and MSE are summarized in Tables 22 and 23 respectively.

Table 22: Estimation MBE of $\hat{\sigma^2_\epsilon}_{OLS}$ and $\hat{\sigma^2_\epsilon}_{MME}$ under re-centered χ^2-distributed measurement error with different variances σ^2_u and residual error with different variances σ^2_ϵ.

	$\sigma^2_\epsilon = 1.5$		$\sigma^2_\epsilon = 2.0$		$\sigma^2_\epsilon = 2.5$	
	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$
$\sigma^2_u = 1.5$	5.893220	-0.056369	5.925753	-0.027765	5.910017	-0.038214
$\sigma^2_u = 2.0$	7.941942	0.031871	7.883551	-0.018494	7.897804	-0.016332
$\sigma^2_u = 2.5$	9.786872	-0.078416	9.747000	-0.111938	9.862407	-0.006246
$\sigma^2_u = 3.0$	11.764364	-0.036215	11.771377	-0.019330	11.577529	-0.215025
$\sigma^2_u = 4.5$	17.515282	-0.048596	17.354157	-0.191983	17.339678	-0.183992
$\sigma^2_u = 6.0$	23.076080	-0.143101	23.001147	-0.215374	22.952380	-0.258514
$\sigma^2_u = 8.0$	30.188767	-0.377915	30.370164	-0.198146	30.555481	-0.029103
$\sigma^2_u = 10.0$	37.399951	-0.459851	37.362384	-0.572358	37.111804	-0.761288

We again have similar observations to the normal and Student's t cases: The MBE of $\hat{\sigma^2_\epsilon}_{OLS}$ is positive and increasing in σ^2_u. The MBE of $\hat{\sigma^2_\epsilon}_{MME}$ is negative and seems to be decreasing in σ^2_u, but the decreasing trend is not that clear. For MSE, the MSE of both $\hat{\sigma^2_\epsilon}_{OLS}$ and $\hat{\sigma^2_\epsilon}_{MME}$ are both increasing in σ^2_u, and the MSE of $\hat{\sigma^2_\epsilon}_{MME}$ is smaller than that of $\hat{\sigma^2_\epsilon}_{OLS}$.

The MBE and MSE results after taking difference with the normal case are shown in Tables 24 and 25 respectively.

For the difference in MBE between the χ^2-distributed and normal-distributed cases, we find that the difference for both $\hat{\sigma^2_\epsilon}_{OLS}$ and $\hat{\sigma^2_\epsilon}_{MME}$ are negative and are roughly decreasing. Because $\hat{\sigma^2_\epsilon}_{OLS}$ is negative in both cases, this means that the MBE for $\hat{\sigma^2_\epsilon}_{OLS}$ is closer to zero under χ^2-distributed errors than normal-distributed errors. Likewise, because $\hat{\sigma^2_\epsilon}_{MME}$ is positive in both cases, this

Table 23: Estimation MSE of $\hat{\sigma^2_\epsilon}_{OLS}$ and $\hat{\sigma^2_\epsilon}_{MME}$ under re-centered χ^2-distributed measurement error with different variances σ_u^2 and residual error with different variances σ_ϵ^2.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$
$\sigma_u^2 = 1.5$	57.156508	23.096732	59.199434	24.756849	59.595066	25.384590
$\sigma_u^2 = 2.0$	94.420321	32.593721	94.155136	33.209791	96.123602	35.016405
$\sigma_u^2 = 2.5$	135.176838	41.299873	134.988694	41.911471	142.479071	47.127930
$\sigma_u^2 = 3.0$	187.194973	51.476957	187.669689	52.236293	182.631566	51.413928
$\sigma_u^2 = 4.5$	385.846689	85.801536	378.435062	83.682260	379.322162	85.748375
$\sigma_u^2 = 6.0$	638.422049	118.133857	636.923293	119.730214	636.076559	122.171158
$\sigma_u^2 = 8.0$	1057.790558	169.895222	1077.712211	178.676540	1093.595726	183.465213
$\sigma_u^2 = 10.0$	1597.776001	237.689528	1599.654260	240.683626	1580.838713	245.632506

Table 24: Difference between re-centered χ^2-distributed error case and normal-distributed error case: Estimation MBE of $\hat{\sigma^2_\epsilon}_{OLS}$ and $\hat{\sigma^2_\epsilon}_{MME}$ under different σ_u^2 and σ_ϵ^2.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$
$\sigma_u^2 = 1.5$	-0.007185	-0.006136	-0.029600	-0.039971	-0.037000	-0.039569
$\sigma_u^2 = 2.0$	0.052979	0.050799	-0.017770	-0.007594	0.018494	0.009332
$\sigma_u^2 = 2.5$	0.002576	-0.009790	-0.055717	-0.040136	-0.006909	-0.019097
$\sigma_u^2 = 3.0$	-0.033836	-0.033334	0.014211	0.021778	-0.215645	-0.212096
$\sigma_u^2 = 4.5$	-0.082670	-0.096197	-0.112445	-0.129886	-0.115736	-0.102601
$\sigma_u^2 = 6.0$	-0.044116	-0.079320	-0.032447	-0.086333	-0.180013	-0.175376
$\sigma_u^2 = 8.0$	-0.217947	-0.227648	0.150550	0.144760	0.165770	0.147509
$\sigma_u^2 = 10.0$	-0.247564	-0.362826	-0.361667	-0.494397	-0.396595	-0.450374

shows that the MBE for $\hat{\sigma^2_\epsilon}_{MME}$ is farther away from zero under χ^2-distributed errors than normal-distributed errors.

For the difference in MSE, we observe that the difference is always positive, meaning that the MSE under χ^2-distributed errors is greater than that under normal-distributed errors. This difference is increasing in σ_u^2, and the difference when using $\hat{\sigma^2_\epsilon}_{MME}$ is greater than that for $\hat{\sigma^2_\epsilon}_{OLS}$.

Table 25: Difference between re-centered χ^2-distributed error case and normal-distributed error case: Estimation MBE of $\hat{\sigma^2_\epsilon}_{OLS}$ and $\hat{\sigma^2_\epsilon}_{MME}$ under different σ^2_u and σ^2_ϵ.

	$\sigma^2_\epsilon = 1.5$		$\sigma^2_\epsilon = 2.0$		$\sigma^2_\epsilon = 2.5$	
	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$	$\hat{\sigma^2_\epsilon}_{OLS}$	$\hat{\sigma^2_\epsilon}_{MME}$
$\sigma^2_u = 1.5$	18.625732	19.316444	19.317283	20.285012	19.385832	20.449473
$\sigma^2_u = 2.0$	26.080697	26.352261	24.935389	26.299499	26.589421	27.377628
$\sigma^2_u = 2.5$	30.552304	32.194016	29.232236	32.005644	34.424333	36.166174
$\sigma^2_u = 3.0$	35.733276	38.742914	36.149158	38.614920	28.888130	36.250397
$\sigma^2_u = 4.5$	50.805596	59.223162	47.415921	56.284041	47.500987	57.305102
$\sigma^2_u = 6.0$	61.591645	72.307314	61.724976	71.979267	54.690090	72.954298
$\sigma^2_u = 8.0$	61.959650	90.747114	93.201731	99.368226	94.810411	100.064881
$\sigma^2_u = 10.0$	76.785113	119.714154	70.779297	120.537955	63.864480	120.335609

5 Conclusion

We have presented some numerical results on the case of non-normality under the errors-in-variables model, and have shown how non-normality in the measurement error affects the estimation of the regression parameters β and σ^2_ϵ. We have compared the errors in different metrics when using different distributions and different estimators for β and σ^2_ϵ.

Future work can be extended upon this topic. We have only presented the numerical results arising from problem of estimation, and the problem of inference of statistical parameters has not been explored in this work. Exact inference of statistical parameters under non-normality can also be explored. The effects of non-normality on multivariate errors-in-variables models and other estimators can also be explored in the future.

6 Acknowledgments

I would like to thank Dr. Raymond W.L. Wong for his support and guidance for this directed studies. I would also like to thank the HKU Department of Statistics and Actuarial Science for providing me with this opportunity to pursue in this directed studies capstone project.

7 Appendix

The computer code used in this project can be found at `https://github.com/jevrii/fyp`.

```
import numpy as np
import matplotlib.pyplot as plt

iterations = 10000
n = 30
secret_beta = 2

var_u_list = [0, 1.5, 2, 2.5,3.0,4.5,6.0,8.0,10.0]
var_eps_list = [1.5, 2, 2.5]

var_u_list = [0, 1.5, 2, 2.5,3.0,4.5,6.0,8.0,10.0]
var_eps_list = [1.5, 2, 2.5]

def experiment(distribution_label, gen_err_u, gen_err_eps):
    print(f"{distribution_label}␣distributed␣epsilon")

    beta_dict = {}
    sigma_dict = {}

    for var_u in var_u_list:
        for var_eps in var_eps_list:
            bias_beta_ols = np.array([])
            sqerr_beta_ols = np.array([])
            bias_beta_mme = np.array([])
            sqerr_beta_mme = np.array([])

            bias_sigma_ols = np.array([])
            sqerr_sigma_ols = np.array([])
            bias_sigma_mme = np.array([])
            sqerr_sigma_mme = np.array([])

            for _ in range(iterations):
                x = np.linspace(-20,20,n)
                if var_u > 0:
                    x_obs = x + gen_err_u(var_u, n)
                else:
                    x_obs = x
                y = x*secret_beta + gen_err_x(var_eps, n)

                beta_ols = np.cov(x_obs, y, ddof=0)[0][1]/np.var(
```

```
        ↪ x_obs, ddof=0)
    beta_mme = np.cov(x_obs, y, ddof=1)[0][1]/(np.var(
        ↪ x_obs, ddof=1) - var_u)

    vx = np.var(x_obs, ddof=0) - var_u
    vu = var_u
    l = vx/(vx+vu);

    sigma_ols = np.sum((y - beta_ols * x_obs) ** 2) / (
        ↪ n-1)
    sigma_mme = np.sum((y - beta_ols * x_obs) ** 2) / (
        ↪ n-1) - (1-l)**2*beta_mme**2*vx - l**2*
        ↪ beta_mme**2*vu

    bias_beta_ols = np.append(bias_beta_ols, (beta_ols-
        ↪ secret_beta))
    sqerr_beta_ols = np.append(sqerr_beta_ols, (
        ↪ beta_ols-secret_beta)**2)
    bias_beta_mme = np.append(bias_beta_mme, (beta_mme-
        ↪ secret_beta))
    sqerr_beta_mme = np.append(sqerr_beta_mme, (
        ↪ beta_mme-secret_beta)**2)

    bias_sigma_ols = np.append(bias_sigma_ols,
        ↪ sigma_ols-var_eps)
    sqerr_sigma_ols = np.append(sqerr_sigma_ols, (
        ↪ sigma_ols-var_eps)**2)
    bias_sigma_mme = np.append(bias_sigma_mme,
        ↪ sigma_mme-var_eps)
    sqerr_sigma_mme = np.append(sqerr_sigma_mme, (
        ↪ sigma_mme-var_eps)**2)

print("var_u=%.2f,␣var_eps=%.2f:␣␣OLS:␣MBE␣=␣%f,␣MSE␣=
    ↪ ␣%f;␣Sigma␣MBE␣=␣%f,␣MSE␣=␣%f"%(var_u, var_eps,
    ↪ bias_beta_ols.mean(), sqerr_beta_ols.mean(),
    ↪ bias_sigma_ols.mean(), sqerr_sigma_ols.mean()))

print("var_u=%.2f,␣var_eps=%.2f:␣␣MME:␣MBE␣=␣%f,␣MSE␣=
    ↪ ␣%f;␣Sigma␣MBE␣=␣%f,␣MSE␣=␣%f"%(var_u, var_eps,
    ↪ bias_beta_mme.mean(), sqerr_beta_mme.mean(),
    ↪ bias_sigma_mme.mean(), sqerr_sigma_mme.mean()))

beta_dict[(var_u, var_eps)] = {}
beta_dict[(var_u, var_eps)]['bias_ols'] =
    ↪ bias_beta_ols.mean()
beta_dict[(var_u, var_eps)]['sqerr_ols'] =
```

```
            ↪ sqerr_beta_ols.mean()
        beta_dict[(var_u, var_eps)]['bias_mme'] =
            ↪ bias_beta_mme.mean()
        beta_dict[(var_u, var_eps)]['sqerr_mme'] =
            ↪ sqerr_beta_mme.mean()

        sigma_dict[(var_u, var_eps)] = {}
        sigma_dict[(var_u, var_eps)]['bias_ols'] =
            ↪ bias_sigma_ols.mean()
        sigma_dict[(var_u, var_eps)]['sqerr_ols'] =
            ↪ sqerr_sigma_ols.mean()
        sigma_dict[(var_u, var_eps)]['bias_mme'] =
            ↪ bias_sigma_mme.mean()
        sigma_dict[(var_u, var_eps)]['sqerr_mme'] =
            ↪ sqerr_sigma_mme.mean()


    return beta_dict, sigma_dict

# experiment under normal-distributed errors

def gen_err_x(var, n):
    return np.random.normal(scale=np.sqrt(var), size=n)

def gen_err_eps(var, n):
    return np.random.normal(scale=np.sqrt(var), size=n)

beta_dict_normal, sigma_dict_normal = experiment("Normal",
    ↪ gen_err_x, gen_err_eps)

# experiment under Student's t-distributed errors

def gen_err_x(var, n):
    return np.random.standard_t(df=2*var/(var-1), size=n) # var=df
        ↪ /(df-2)

def gen_err_eps(var, n):
    return np.random.standard_t(df=2*var/(var-1), size=n)

beta_dict_t, sigma_dict_t = experiment("Student-T", gen_err_x,
    ↪ gen_err_eps)

# experiment under re-centered Chi^2-distributed errors

def gen_err_x(var, n):
    return np.random.chisquare(df=var/2.0, size=n) - var/2.0# var
```

```
    ↪ =2*dof

def gen_err_eps(var, n):
    return np.random.chisquare(df=var/2.0, size=n) - var/2.0

beta_dict_chi, sigma_dict_chi = experiment("ChiSq", gen_err_x,
    ↪ gen_err_eps)
```

References

[1] Pischke S. Lecture Notes on Measurement Error. 2007.

[2] Gillard J. Method of Moments Estimation in Linear Regression with Errors in both Variables. Communications in Statistics - Theory and Methods. 2014;43(15):3208-22.