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 observation error, and that the errors are normally distributed. However, this may not be the case in real-life applications, where observation 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 empirical results from simulations based on normal, Student's t and χ^2 distributions on the estimation of regression parameters β and 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

1

$(\tilde{x}_i, \tilde{y}_i)$, 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 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 observation matter u. Thus, there is a need to investigate the impacts of non-normality on the *errors-in-variables model*.

In this project, we investigate how the non-normality of errors in both the explanatory variable x and the residual of the dependent variable y affects the estimation of the regression coefficient $\hat{\beta}$ 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 computer simulation technique to produce results. Finally, we will present the results and findings from the computer simulations. The code used in this project 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:

$$y_i = \beta(\tilde{x}_i - u_i) + \epsilon_i = \beta\tilde{x} + (\epsilon - \beta u) \tag{2}$$

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*. Unlike the case of with 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. n order to obtain unbiased and consistent estimates, we would have to resort to the method of moments (MME) estimators instead.

2.1   OLS and MME for β

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

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

Because ϵ, u and x are independent to each other, we can obtain

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

where

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

This λ is also called the reliability 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 the OLS $\hat{\beta}$ as $\sigma_{\tilde{x}}^2 = \sigma_x^2 + \sigma_u^2$, and $\sigma_{\tilde{x}}^2$ can be directly measured from the observed data.

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

Table 1: Various MME estimators for β

Assumption	Method of Moments Estimator
σ_x^2 known	$\frac{cov(\tilde{x},y)}{\sigma_x^2}$
σ_u^2 known	$\frac{cov(\tilde{x},y)}{var(\tilde{x})-\sigma_u^2}$
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 obtian the value of σ_u^2, 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_ϵ^2

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

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

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

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

$$
\text{plim } \hat{\sigma_\epsilon^2} = \sigma_\epsilon^2 + (1-\lambda)^2\beta^2\sigma_x^2 + \lambda^2\beta^2\sigma_u^2
\tag{7}
$$

In order to obtain the MME estimator for σ_ϵ^2, similar to the case for $\hat{\beta}_{MME}$, if we have some prior knowledge on the observation error, we can obtain consistent

estimates for σ_x^2, σ_u, λ and β. Thus, rearranging the expression for $\hat{\sigma}_\epsilon^2$ gives:

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

Based on the relation $\tilde{x}_i = x_i + u_i$, we can obtain $var(x) = var(\tilde{x}) - var(u)$. Thus, if σ_u^2 is known, and we use $var(x)$ as an estimator for σ_x^2, we can substitute into 8:

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

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}, xhi. 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 observation error u, we choose a distribution where the observation errors u_i are sampled from. The model parameters of this distribution is 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 model parameters of this distribution is 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 draw independently n instances of explanatory variable x_i from the uniform distribution defined by the bounds x_{lo} and x_{hi}, $U(x_{lo}, x_{hi})$. Afterwards, for each of the instances x_i, we compute $y_i = \beta_{truth} \cdot x_i + \epsilon_i$, where 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 observation error u_i is added for each x_i. u_i is drawn independently from the distribution for observation 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 observation error σ_u^2 is known. The relevant expressions for the estimators are listed below:

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

$$\hat{\beta}_{MME} = \frac{var(\tilde{x})}{var(\tilde{x}) - \sigma_u^2} \qquad (10)$$

$$\hat{\sigma^2}_{\epsilon\,OLS} = var\left(y - \hat{\beta}_{OLS}\tilde{x}\right)$$

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

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 observation 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 (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 observation 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. 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 is known to have heavier tails than the normal distribution, and the χ^2 distribution is known to be skewed. They pose substantial differences from the normal distribution which has light tails and is symmetric at 0.

We will then perform the simulations using these distributions as the underlying sampling distributions of the observation error u and residual error ϵ, as

---

Procedure 1: Computer simulation procedure

Initialize β_{truth}, n, x_{lo}, x_{hi};

forall $\sigma_u^2 \in \{list\ of\ observation\ variance\}$ **do**

    forall $\sigma_\epsilon^2 \in \{list\ of\ residual\ variance\}$ **do**

        Initialize $MSE_{\hat{\beta}_{OLS}} = 0$, $MSE_{\hat{\beta}_{MME}} = 0$, $MBE_{\hat{\beta}_{OLS}} = 0$, $MBE_{\hat{\beta}_{MME}} = 0.$;

        Initialize $MSE_{\hat{\sigma}^2_{\epsilon\,OLS}} = 0$, $MSE_{\hat{\sigma}^2_{\epsilon\,MME}} = 0$, $MBE_{\hat{\sigma}^2_{\epsilon\,OLS}} = 0$, $MBE_{\hat{\sigma}^2_{\epsilon\,MME}} = 0$;

        for *10000 iterations* **do**

            Draw independently n instances of explanatory variable x_i from $U(x_{lo}, x_{hi})$;

            Compute n observations of $y_i = \beta_{truth} \cdot x_i + \epsilon_i$, where ϵ_i is drawn independently from a distribution with variance σ_ϵ^2;

            Compute n instances of observed explanatory variable $\tilde{x}_i = x_i + u_i$, where observation u_i is drawn independently from a distribution with variance σ_u^2;

            Calculate $\hat{\beta}_{OLS}$, $\hat{\beta}_{MME}$, $\hat{\sigma}^2_{\epsilon\,OLS}$, $\hat{\sigma}^2_{\epsilon\,MSE}$ from the observations $\tilde{x}$, y, and prior information on observation error σ_u^2;

            Update $MSE_{\hat{\beta}_{OLS}}$, $MSE_{\hat{\beta}_{MME}}$, $MBE_{\hat{\beta}_{OLS}}$, $MBE_{\hat{\beta}_{MME}}$, $MSE_{\hat{\sigma}^2_{\epsilon\,OLS}}$, $MSE_{\hat{\sigma}^2_{\epsilon\,MME}}$, $MBE_{\hat{\sigma}^2_{\epsilon\,OLS}}$, $MBE_{\hat{\sigma}^2_{\epsilon\,MME}}$;

        end

    end

end

---

described in the previous subsection.

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 observation 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 observation error (i.e. $\sigma_u^2 = 0$). For the cases under the presence of observation 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 observation error, and cases under the presence of observation 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 $(\tilde{x}_i, y_i)$:

$$\hat{\beta}_{OLS} = \frac{cov(\tilde{x}, y)}{var(\tilde{x})} = \frac{cov(x + u, \beta x + \epsilon)}{var(x + u)}$$
$$\hat{\beta}_{MME} = \frac{var(\tilde{x})}{var(\tilde{x}) - \sigma_u^2} \tag{11}$$

4.1.1 Absence of observation error

Under the absence of observation 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
$\sigma_\epsilon^2 = 1.5$	-0.000260	0.000252	0.000052
$\sigma_\epsilon^2 = 2.0$	0.000151	0.000114	-0.000356
$\sigma_\epsilon^2 = 2.5$	-0.000176	-0.000105	0.000112

butions. There is no clear trend of increasing/decreasing pattern of the MBE for each of the distributions. This is mostly because the bias is 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
$\sigma_\epsilon^2 = 1.5$	0.000352	0.000351	0.000343
$\sigma_\epsilon^2 = 2.0$	0.000470	0.000478	0.000458
$\sigma_\epsilon^2 = 2.5$	0.000588	0.000594	0.000573

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. Furthermore, we can observe that the MSE for χ^2 distribution is the least in all 3 experiments, and the MSE for Student's t distribution is the greatest in 2 out of 3 experiments.

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

Under the presence of observation 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 observation error, the OLS estimator is inconsistent but the MME estimator is consistent.

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 $\hat{\beta}_{OLS} = \frac{\beta \sigma_x^2}{\sigma_x^2 + \sigma_u^2} = \lambda \cdot \beta_{truth}$ and is 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 observation error with different variances σ_u^2, and under normal-distributed residual error with different variances σ_ϵ^2.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$
$\sigma_u^2 = 1.5$	-0.019467	0.000711	-0.018743	0.001453	-0.018929	0.001258
$\sigma_u^2 = 2.0$	-0.025206	0.001655	-0.023977	0.002920	-0.025265	0.001602
$\sigma_u^2 = 2.5$	-0.030908	0.002607	-0.030422	0.003106	-0.030886	0.002635
$\sigma_u^2 = 3.0$	-0.036965	0.003140	-0.036627	0.003527	-0.036902	0.003217
$\sigma_u^2 = 4.5$	-0.055003	0.004770	-0.056174	0.003501	-0.056439	0.003171
$\sigma_u^2 = 6.0$	-0.072019	0.007202	-0.071406	0.007854	-0.073418	0.005616
$\sigma_u^2 = 8.0$	-0.096652	0.007750	-0.096022	0.008548	-0.096813	0.007639
$\sigma_u^2 = 10.0$	-0.119142	0.010226	-0.121257	0.007779	-0.117934	0.011878

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

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$
$\sigma_u^2 = 1.5$	0.002049	0.001761	0.002120	0.001862	0.002254	0.001992
$\sigma_u^2 = 2.0$	0.002801	0.002331	0.002747	0.002338	0.003020	0.002552
$\sigma_u^2 = 2.5$	0.003447	0.002739	0.003563	0.002895	0.003810	0.003124
$\sigma_u^2 = 3.0$	0.004320	0.003313	0.004423	0.003452	0.004584	0.003595
$\sigma_u^2 = 4.5$	0.007218	0.004996	0.007379	0.005007	0.007404	0.004989
$\sigma_u^2 = 6.0$	0.010421	0.006639	0.010245	0.006529	0.010711	0.006687
$\sigma_u^2 = 8.0$	0.015781	0.008820	0.015708	0.008875	0.016050	0.009091
$\sigma_u^2 = 10.0$	0.022013	0.011578	0.022421	0.011359	0.021891	0.011826

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 $\hat{\beta}_{MME}$ (roughly) 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 observation error and Student's t-distributed residual error

We perform the similar simulations for Student's t distributed observation error and residual error for different variances. Elementary statistical theory tells us that the variance for a Student's t distribution with $\nu > 0$ 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 observation error with different variances σ_u^2, and under Student's t-distributed residual error with different variances σ_ϵ^2.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$
$\sigma_u^2 = 1.5$	-0.018770	0.001429	-0.019233	0.000955	-0.018866	0.001329
$\sigma_u^2 = 2.0$	-0.024866	0.002008	-0.025077	0.001785	-0.025835	0.001001
$\sigma_u^2 = 2.5$	-0.030725	0.002796	-0.030205	0.003343	-0.030965	0.002552
$\sigma_u^2 = 3.0$	-0.035638	0.004529	-0.035464	0.004732	-0.034195	0.006041
$\sigma_u^2 = 4.5$	-0.047532	0.012747	-0.047863	0.012389	-0.047730	0.012474
$\sigma_u^2 = 6.0$	-0.053997	0.026746	-0.054018	0.026632	-0.055230	0.025393
$\sigma_u^2 = 8.0$	-0.063497	0.044623	-0.062978	0.045236	-0.060738	0.047766
$\sigma_u^2 = 10.0$	-0.066641	0.070167	-0.065806	0.070902	-0.065357	0.071511

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.

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

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$
$\sigma_u^2 = 1.5$	0.002021	0.001761	0.002213	0.001939	0.002321	0.002065
$\sigma_u^2 = 2.0$	0.002921	0.002473	0.003258	0.002809	0.003236	0.002744
$\sigma_u^2 = 2.5$	0.004083	0.003415	0.004288	0.003668	0.004529	0.003870
$\sigma_u^2 = 3.0$	0.005745	0.004929	0.005835	0.005041	0.005557	0.004851
$\sigma_u^2 = 4.5$	0.011736	0.010821	0.011447	0.010442	0.010748	0.009719
$\sigma_u^2 = 6.0$	0.013428	0.013042	0.013088	0.012611	0.014464	0.013965
$\sigma_u^2 = 8.0$	0.018133	0.019265	0.017601	0.018818	0.017577	0.019249
$\sigma_u^2 = 10.0$	0.020794	0.025946	0.019452	0.024433	0.019140	0.024228

However, when comparing $\hat{\beta}_{OLS}$ and $\hat{\beta}_{MME}$, we find that switching from $\hat{\beta}_{OLS}$ to $\hat{\beta}_{MME}$ does not provide that much of a decrease in absolute MBE or MSE 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.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$
$\sigma_u^2 = 1.5$	-0.000028	-0.000000	0.000093	0.000077	0.000067	0.000073
$\sigma_u^2 = 2.0$	0.000120	0.000142	0.000511	0.000471	0.000215	0.000191
$\sigma_u^2 = 2.5$	0.000636	0.000676	0.000725	0.000773	0.000719	0.000746
$\sigma_u^2 = 3.0$	0.001425	0.001616	0.001412	0.001589	0.000973	0.001257
$\sigma_u^2 = 4.5$	0.004518	0.005825	0.004068	0.005434	0.003343	0.004731
$\sigma_u^2 = 6.0$	0.003007	0.006403	0.002843	0.006083	0.003753	0.007279
$\sigma_u^2 = 8.0$	0.002352	0.010444	0.001893	0.009943	0.001527	0.010159
$\sigma_u^2 = 10.0$	-0.001219	0.014368	-0.002969	0.013074	-0.002751	0.012402

It should be noted that when the variance of a Student's t distribution

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.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$
$\sigma_u^2 = 1.5$	0.000697	0.000718	-0.000491	-0.000498	0.000063	0.000072
$\sigma_u^2 = 2.0$	0.000340	0.000354	-0.001099	-0.001135	-0.000570	-0.000601
$\sigma_u^2 = 2.5$	0.000183	0.000189	0.000218	0.000237	-0.000079	-0.000083
$\sigma_u^2 = 3.0$	0.001327	0.001390	0.001162	0.001205	0.002707	0.002824
$\sigma_u^2 = 4.5$	0.007471	0.007977	0.008311	0.008888	0.008708	0.009303
$\sigma_u^2 = 6.0$	0.018022	0.019544	0.017388	0.018778	0.018188	0.019777
$\sigma_u^2 = 8.0$	0.033155	0.036873	0.033045	0.036688	0.036074	0.040126
$\sigma_u^2 = 10.0$	0.052501	0.059941	0.055451	0.063122	0.052577	0.059633

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 Student's t distribution increases, it converges to the normal distribution. Thus, for some low fixed variance, we should expect similar findings 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 increases to a certain point (around at $\sigma_u^2 = 4.5$), then again starts to decrease. As the bias of $\hat{\beta}_{OLS}$ is negative in both Student's t-distributed errors and normal-distributed errors, this shows that $\hat{\beta}_{OLS}$ under Student's t errors is biased downwards less. On the other hand, for $\hat{\beta}_{MME}$, we observe that the difference is increasing throughout as σ_u^2 is increased.

We then focus on the comparisons in MSE. Although the trends for small σ_u^2 is not clear, for both $\hat{\beta}_{OLS}$ and $\hat{\beta}_{MME}$ we observe that the MSE under Student's t-distributed errors is in general greater than that under normal-distributed errors, and that this difference is in general increasing in σ_u^2.

4.1.4 χ^2-distributed observation error

We perform the similar simulations for χ^2 distributed observation 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 $2 \cdot \frac{\delta}{2}$ 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 χ^2-distributed observation error with different variances σ_u^2 and residual error with different variances σ_ϵ^2.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$
$\sigma_u^2 = 1.5$	-0.018714	0.001479	-0.018709	0.001491	-0.018640	0.001558
$\sigma_u^2 = 2.0$	-0.025420	0.001430	-0.024128	0.002761	-0.025012	0.001851
$\sigma_u^2 = 2.5$	-0.030573	0.002949	-0.031135	0.002355	-0.031329	0.002149
$\sigma_u^2 = 3.0$	-0.036298	0.003864	-0.036901	0.003226	-0.037024	0.003085
$\sigma_u^2 = 4.5$	-0.054917	0.004847	-0.055502	0.004227	-0.054086	0.005756
$\sigma_u^2 = 6.0$	-0.072913	0.006221	-0.072631	0.006482	-0.072691	0.006443
$\sigma_u^2 = 8.0$	-0.097304	0.007052	-0.095563	0.009114	-0.097776	0.006451
$\sigma_u^2 = 10.0$	-0.119601	0.009722	-0.119559	0.009724	-0.117377	0.012370

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

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$
$\sigma_u^2 = 1.5$	0.002214	0.001963	0.002375	0.002129	0.002378	0.002131
$\sigma_u^2 = 2.0$	0.003024	0.002552	0.002964	0.002557	0.003111	0.002658
$\sigma_u^2 = 2.5$	0.003698	0.003029	0.003881	0.003179	0.003957	0.003242
$\sigma_u^2 = 3.0$	0.004428	0.003481	0.004697	0.003716	0.004774	0.003782
$\sigma_u^2 = 4.5$	0.007516	0.005330	0.007542	0.005260	0.007606	0.005527
$\sigma_u^2 = 6.0$	0.011204	0.007381	0.011033	0.007194	0.011208	0.007384
$\sigma_u^2 = 8.0$	0.016869	0.009983	0.016494	0.009954	0.016978	0.009951
$\sigma_u^2 = 10.0$	0.022575	0.012067	0.022936	0.012550	0.022595	0.012841

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.

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.

When comparing findings from χ^2-distributed errors and normal-distributed errors, it should be noted that the χ^2 distribution re-centered at mean zero

Table 12: Difference between 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.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$
$\sigma_u^2 = 1.5$	0.000752	0.000768	0.000034	0.000038	0.000289	0.000301
$\sigma_u^2 = 2.0$	-0.000214	-0.000225	-0.000150	-0.000159	0.000253	0.000249
$\sigma_u^2 = 2.5$	0.000335	0.000342	-0.000712	-0.000751	-0.000442	-0.000486
$\sigma_u^2 = 3.0$	0.000667	0.000724	-0.000274	-0.000301	-0.000122	-0.000133
$\sigma_u^2 = 4.5$	0.000086	0.000077	0.000672	0.000725	0.002353	0.002585
$\sigma_u^2 = 6.0$	-0.000895	-0.000981	-0.001225	-0.001371	0.000728	0.000827
$\sigma_u^2 = 8.0$	-0.000651	-0.000698	0.000460	0.000566	-0.000963	-0.001188
$\sigma_u^2 = 10.0$	-0.000459	-0.000504	0.001698	0.001944	0.000557	0.000492

Table 13: Difference between 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.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$	$\hat{\beta}_{OLS}$	$\hat{\beta}_{MME}$
$\sigma_u^2 = 1.5$	0.000165	0.000202	0.000255	0.000267	0.000124	0.000139
$\sigma_u^2 = 2.0$	0.000223	0.000221	0.000216	0.000219	0.000090	0.000105
$\sigma_u^2 = 2.5$	0.000251	0.000290	0.000318	0.000285	0.000147	0.000118
$\sigma_u^2 = 3.0$	0.000108	0.000167	0.000274	0.000264	0.000190	0.000188
$\sigma_u^2 = 4.5$	0.000298	0.000334	0.000164	0.000253	0.000202	0.000538
$\sigma_u^2 = 6.0$	0.000783	0.000742	0.000788	0.000666	0.000498	0.000697
$\sigma_u^2 = 8.0$	0.001088	0.001163	0.000785	0.001079	0.000928	0.000860
$\sigma_u^2 = 10.0$	0.000562	0.000489	0.000516	0.001191	0.000705	0.001015

converges to a normal distribution with the same variance as 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 $(\tilde{x}_i, y_i)$:

$$\hat{\sigma}^2_{\epsilon\,OLS} = var\left(y - \hat{\beta}_{OLS}\tilde{x}\right)$$
$$\hat{\sigma}^2_{\epsilon\,MME} = \hat{\sigma}^2_{\epsilon\,OLS} - (1-\lambda)^2\hat{\beta}^2_{MME}(var(\tilde{x}) - \sigma_u^2) - \lambda^2\hat{\beta}^2_{MME} \tag{12}$$

4.2.1 Absence of observation error

Under the absence of observation error, the OLS and MME estimators of σ_ϵ^2 are the same.

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
$\sigma_\epsilon^2 = 1.5$	0.000038	0.006738	-0.009608
$\sigma_\epsilon^2 = 2.0$	0.000135	-0.010998	-0.005749
$\sigma_\epsilon^2 = 2.5$	-0.001355	0.018754	0.026852

Table 14 summarizes the MBE of σ_ϵ^2 when $\sigma_u^2 = 0$ for different error distributions. 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.

TODO: MSE

4.2.2 Normal-distributed observation error

We again perform the estimation of σ_ϵ^2 under normal-distributed observation error and normal-distributed residual error, both using different variances. The simulated MBE and MSE are tabulated in Tables 15 and **??** respectively.

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 $\hat{\sigma}^2_{\epsilon\,OLS}$.

Table 15: Estimation MBE of $\hat{\sigma}^2_{\epsilon\,OLS}$ and $\hat{\sigma}^2_{\epsilon\,MME}$ under normal-distributed observation error with different variances σ^2_u and residual error with different variances σ^2_ϵ.

	$\sigma^2_\epsilon = 1.5$		$\sigma^2_\epsilon = 2.0$		$\sigma^2_\epsilon = 2.5$	
	$\hat{\sigma}^2_{\epsilon\,OLS}$	$\hat{\sigma}^2_{\epsilon\,MME}$	$\hat{\sigma}^2_{\epsilon\,OLS}$	$\hat{\sigma}^2_{\epsilon\,MME}$	$\hat{\sigma}^2_{\epsilon\,OLS}$	$\hat{\sigma}^2_{\epsilon\,MME}$
$\sigma^2_u = 1.5$	5.944546	-0.005842	5.959183	0.010743	5.931388	-0.014544
$\sigma^2_u = 2.0$	7.886910	-0.024392	7.899105	-0.010766	7.826871	-0.081642
$\sigma^2_u = 2.5$	9.825091	-0.041342	9.837348	-0.020427	9.810995	-0.038816
$\sigma^2_u = 3.0$	11.752766	-0.042576	11.827087	0.014613	11.813583	0.010360
$\sigma^2_u = 4.5$	17.563742	-0.015657	17.409776	-0.158051	17.471703	-0.083689
$\sigma^2_u = 6.0$	23.104857	-0.118361	23.084622	-0.121116	23.031219	-0.217114
$\sigma^2_u = 8.0$	30.385471	-0.234288	30.443624	-0.211370	30.481465	-0.197597
$\sigma^2_u = 10.0$	37.582805	-0.319186	37.466108	-0.471654	37.645894	-0.250725

Table 16: Estimation MSE of $\hat{\sigma}^2_{\epsilon\,OLS}$ and $\hat{\sigma}^2_{\epsilon\,MME}$ under normal-distributed observation error with different variances σ^2_u and residual error with different variances σ^2_ϵ.

	$\sigma^2_\epsilon = 1.5$		$\sigma^2_\epsilon = 2.0$		$\sigma^2_\epsilon = 2.5$	
	$\hat{\sigma}^2_{\epsilon\,OLS}$	$\hat{\sigma}^2_{\epsilon\,MME}$	$\hat{\sigma}^2_{\epsilon\,OLS}$	$\hat{\sigma}^2_{\epsilon\,MME}$	$\hat{\sigma}^2_{\epsilon\,OLS}$	$\hat{\sigma}^2_{\epsilon\,MME}$
$\sigma^2_u = 1.5$	39.106310	3.818130	39.828669	4.367114	40.113631	4.986267
$\sigma^2_u = 2.0$	68.251852	6.135680	69.143363	6.878215	68.728170	7.618354
$\sigma^2_u = 2.5$	105.404358	9.135046	106.313434	9.829482	106.637167	10.601876
$\sigma^2_u = 3.0$	149.786174	11.983789	153.149474	13.570971	153.738092	14.681478
$\sigma^2_u = 4.5$	334.039435	26.905638	328.920230	26.949451	333.116014	29.293710
$\sigma^2_u = 6.0$	575.168146	44.066269	576.296059	46.827447	575.827556	48.625474
$\sigma^2_u = 8.0$	993.932566	78.364319	999.708303	80.049208	1003.433290	81.952748
$\sigma^2_u = 10.0$	1517.354461	120.073998	1511.381767	122.354870	1528.536252	126.664946

On the other hand, for $\hat{\sigma}^2_{\epsilon\,MME}$, the MBE is negative and roughly decreasing with σ^2_u. There is no clear trend with σ^2_ϵ.

4.2.3 Student's t-distributed observation error

We perform the similar simulations and estimation of σ^2_ϵ under the case of Student's t-distributed observation error and residual error. The MBE and MSE results are tabulated in Tables 17 and 18 respectively:

We find that the observations are similar to the case of estimating β under Student's t-distributed errors: We find that the 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.

Table 17: Estimation MBE of $\hat{\sigma}^2_{\epsilon\,OLS}$ and $\hat{\sigma}^2_{\epsilon\,MME}$ under normal-distributed observation error with different variances σ^2_u and residual error with different variances σ^2_ϵ.

	$\sigma^2_\epsilon = 1.5$		$\sigma^2_\epsilon = 2.0$		$\sigma^2_\epsilon = 2.5$	
	$\hat{\sigma}^2_{\epsilon\,OLS}$	$\hat{\sigma}^2_{\epsilon\,MME}$	$\hat{\sigma}^2_{\epsilon\,OLS}$	$\hat{\sigma}^2_{\epsilon\,MME}$	$\hat{\sigma}^2_{\epsilon\,OLS}$	$\hat{\sigma}^2_{\epsilon\,MME}$
$\sigma^2_u = 1.5$	5.901136	-0.051106	5.949806	0.002570	5.921078	-0.029820
$\sigma^2_u = 2.0$	7.808126	-0.103737	7.873660	-0.038787	7.785658	-0.128063
$\sigma^2_u = 2.5$	9.403487	-0.473446	9.674954	-0.187735	9.548231	-0.327451
$\sigma^2_u = 3.0$	10.991609	-0.841798	11.224290	-0.613500	11.263660	-0.552675
$\sigma^2_u = 4.5$	14.854718	-2.893005	14.739074	-3.008623	15.304472	-2.413692
$\sigma^2_u = 6.0$	16.737684	-6.991530	16.764042	-6.956211	17.054863	-6.659021
$\sigma^2_u = 8.0$	19.173947	-12.709664	19.076132	-12.745150	19.562751	-12.287313
$\sigma^2_u = 10.0$	21.736085	-18.363607	20.925550	-19.274530	20.519368	-19.725613

Table 18: Estimation MSE of $\hat{\sigma}^2_{\epsilon\,OLS}$ and $\hat{\sigma}^2_{\epsilon\,MME}$ under normal-distributed observation error with different variances σ^2_u and residual error with different variances σ^2_ϵ.

	$\sigma^2_\epsilon = 1.5$		$\sigma^2_\epsilon = 2.0$		$\sigma^2_\epsilon = 2.5$	
	$\hat{\sigma}^2_{\epsilon\,OLS}$	$\hat{\sigma}^2_{\epsilon\,MME}$	$\hat{\sigma}^2_{\epsilon\,OLS}$	$\hat{\sigma}^2_{\epsilon\,MME}$	$\hat{\sigma}^2_{\epsilon\,OLS}$	$\hat{\sigma}^2_{\epsilon\,MME}$
$\sigma^2_u = 1.5$	42.007008	7.330577	44.265186	9.028256	57.591174	22.712470
$\sigma^2_u = 2.0$	85.073222	25.150626	90.945333	30.167468	101.248433	41.912878
$\sigma^2_u = 2.5$	151.504695	66.572613	178.829111	90.060470	216.508799	130.531705
$\sigma^2_u = 3.0$	242.209257	130.228570	296.302408	181.223069	306.113309	190.717377
$\sigma^2_u = 4.5$	674.176226	504.183618	647.724335	480.312491	853.425135	679.140708
$\sigma^2_u = 6.0$	812.257778	650.836042	853.904589	694.262448	961.547494	800.543447
$\sigma^2_u = 8.0$	1213.885175	1159.121657	1260.906736	1218.715769	1323.105685	1251.042547
$\sigma^2_u = 10.0$	1873.268174	2046.381778	1691.114328	1887.541448	1558.572354	1770.970898

Similar to the case of estimating β, we also find that when σ^2_u increases, the improvement of using $\hat{\sigma}^2_{\epsilon\,MME}$ instead of $\hat{\sigma}^2_{\epsilon\,OLS}$ decreases. In fact, we see that when $\sigma^2_u = 10.0$, 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 19 and 20 respectively:

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^2_u. 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

Table 19: 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^2_u and σ^2_ϵ.

	$\sigma^2_\epsilon = 1.5$		$\sigma^2_\epsilon = 2.0$		$\sigma^2_\epsilon = 2.5$	
	$\hat{\sigma}^2_{\epsilon OLS}$	$\hat{\sigma}^2_{\epsilon MME}$	$\hat{\sigma}^2_{\epsilon OLS}$	$\hat{\sigma}^2_{\epsilon MME}$	$\hat{\sigma}^2_{\epsilon OLS}$	$\hat{\sigma}^2_{\epsilon MME}$
$\sigma^2_u = 1.5$	-0.066959	-0.069825	0.001343	0.002872	-0.000262	0.000937
$\sigma^2_u = 2.0$	-0.109399	-0.114493	0.016825	0.012610	-0.150998	-0.158049
$\sigma^2_u = 2.5$	-0.430338	-0.443487	-0.183308	-0.187789	-0.298113	-0.310561
$\sigma^2_u = 3.0$	-0.803268	-0.836666	-0.559063	-0.595413	-0.515155	-0.526378
$\sigma^2_u = 4.5$	-2.704054	-2.880683	-2.761439	-2.954847	-2.155957	-2.311831
$\sigma^2_u = 6.0$	-6.380179	-6.874139	-6.318024	-6.798425	-5.968561	-6.496184
$\sigma^2_u = 8.0$	-11.271448	-12.496345	-11.273698	-12.442319	-10.695063	-11.949910
$\sigma^2_u = 10.0$	-15.743113	-17.890013	-16.732464	-19.050908	-16.942578	-19.309470

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^2_u and σ^2_ϵ.

	$\sigma^2_\epsilon = 1.5$		$\sigma^2_\epsilon = 2.0$		$\sigma^2_\epsilon = 2.5$	
	$\hat{\sigma}^2_{\epsilon OLS}$	$\hat{\sigma}^2_{\epsilon MME}$	$\hat{\sigma}^2_{\epsilon OLS}$	$\hat{\sigma}^2_{\epsilon MME}$	$\hat{\sigma}^2_{\epsilon OLS}$	$\hat{\sigma}^2_{\epsilon MME}$
$\sigma^2_u = 1.5$	2.567444	3.439570	4.459091	4.544724	17.715820	17.829044
$\sigma^2_u = 2.0$	16.250164	18.866020	22.518222	23.336357	30.675791	34.186680
$\sigma^2_u = 2.5$	45.666405	57.182168	71.846646	79.995463	109.022235	119.665395
$\sigma^2_u = 3.0$	90.881267	117.709726	144.389898	167.697770	153.319059	176.251063
$\sigma^2_u = 4.5$	340.965781	477.641574	315.979763	453.268744	521.592655	650.493605
$\sigma^2_u = 6.0$	236.006213	605.414741	277.237814	647.609167	385.797868	751.716660
$\sigma^2_u = 8.0$	217.295838	1082.203769	267.746928	1137.929792	334.182042	1169.656013
$\sigma^2_u = 10.0$	364.077884	1927.985594	164.195051	1764.231817	45.362629	1644.894215

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.

For MSE, we observe that the difference is positive and increasing in σ^2_u. 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 observation 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 21 and 22 respectively.

Table 21: Estimation MBE of $\hat{\sigma}^2_{\epsilon OLS}$ and $\hat{\sigma}^2_{\epsilon MME}$ under χ^2-distributed observation error with different variances σ^2_u and residual error with different variances σ^2_ϵ.

	$\sigma^2_\epsilon = 1.5$		$\sigma^2_\epsilon = 2.0$		$\sigma^2_\epsilon = 2.5$	
	$\hat{\sigma}^2_{\epsilon OLS}$	$\hat{\sigma}^2_{\epsilon MME}$	$\hat{\sigma}^2_{\epsilon OLS}$	$\hat{\sigma}^2_{\epsilon MME}$	$\hat{\sigma}^2_{\epsilon OLS}$	$\hat{\sigma}^2_{\epsilon MME}$
$\sigma^2_u = 1.5$	5.846227	-0.107955	5.852595	-0.097985	5.956031	0.001673
$\sigma^2_u = 2.0$	7.916757	0.010683	7.853233	-0.059238	7.983631	0.070941
$\sigma^2_u = 2.5$	9.848993	-0.016529	9.828657	-0.039154	9.799226	-0.076836
$\sigma^2_u = 3.0$	11.705806	-0.108001	11.658828	-0.148038	11.752845	-0.055206
$\sigma^2_u = 4.5$	17.374873	-0.210306	17.373310	-0.183002	17.402963	-0.193137
$\sigma^2_u = 6.0$	22.814673	-0.435660	22.796974	-0.457350	23.039812	-0.174432
$\sigma^2_u = 8.0$	30.256286	-0.377694	30.314447	-0.328371	30.343469	-0.340969
$\sigma^2_u = 10.0$	37.120092	-0.841162	37.539161	-0.432518	37.117431	-0.840077

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

	$\sigma^2_\epsilon = 1.5$		$\sigma^2_\epsilon = 2.0$		$\sigma^2_\epsilon = 2.5$	
	$\hat{\sigma}^2_{\epsilon OLS}$	$\hat{\sigma}^2_{\epsilon MME}$	$\hat{\sigma}^2_{\epsilon OLS}$	$\hat{\sigma}^2_{\epsilon MME}$	$\hat{\sigma}^2_{\epsilon OLS}$	$\hat{\sigma}^2_{\epsilon MME}$
$\sigma^2_u = 1.5$	56.563197	22.996346	57.309948	23.677762	60.483599	25.715861
$\sigma^2_u = 2.0$	93.680355	32.217787	91.819179	31.360443	98.074845	35.627879
$\sigma^2_u = 2.5$	137.435981	42.521144	137.176570	42.563518	136.832234	42.626278
$\sigma^2_u = 3.0$	184.303172	50.044663	183.957501	50.917001	188.524146	53.241392
$\sigma^2_u = 4.5$	377.787401	82.590910	379.565750	84.655555	381.000349	84.963603
$\sigma^2_u = 6.0$	625.109020	116.664481	623.575613	117.049269	640.975827	123.237757
$\sigma^2_u = 8.0$	1067.198766	175.652108	1074.095498	179.207401	1072.267645	174.251989
$\sigma^2_u = 10.0$	1571.092644	232.458616	1612.083101	239.745418	1581.926597	242.492254

We again observe similar properties to the normal and Student's t cases: For MBE, the MBE of $\hat{\sigma}^2_{\epsilon OLS}$ is positive and increasing in σ^2_u, and that of $\hat{\sigma}^2_{\epsilon MME}$ is negative and decreasing in σ^2_u. 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 23 and 24 respectively.

For the difference in MBE between the χ^2-distributed and normal-distributed cases, we do not see a clear monotonic trend in the difference value, and the sign fluctuates. 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^2_u,

Table 23: Difference between 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.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\sigma^2}_{\epsilon OLS}$	$\hat{\sigma^2}_{\epsilon MME}$	$\hat{\sigma^2}_{\epsilon OLS}$	$\hat{\sigma^2}_{\epsilon MME}$	$\hat{\sigma^2}_{\epsilon OLS}$	$\hat{\sigma^2}_{\epsilon MME}$
$\sigma_u^2 = 1.5$	-0.121868	-0.126673	-0.095868	-0.097683	0.034691	0.032430
$\sigma_u^2 = 2.0$	-0.000769	-0.000073	-0.003602	-0.007841	0.046975	0.040955
$\sigma_u^2 = 2.5$	0.015169	0.013430	-0.029604	-0.039207	-0.047118	-0.059946
$\sigma_u^2 = 3.0$	-0.089071	-0.102870	-0.124525	-0.129951	-0.025971	-0.028909
$\sigma_u^2 = 4.5$	-0.183899	-0.197984	-0.127204	-0.129226	-0.057467	-0.091276
$\sigma_u^2 = 6.0$	-0.303190	-0.318269	-0.285091	-0.299563	0.016387	-0.011594
$\sigma_u^2 = 8.0$	-0.189109	-0.164376	-0.035383	-0.025541	0.085655	-0.003566
$\sigma_u^2 = 10.0$	-0.359106	-0.367568	-0.118853	-0.208896	-0.344515	-0.423934

Table 24: Difference between 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.

	$\sigma_\epsilon^2 = 1.5$		$\sigma_\epsilon^2 = 2.0$		$\sigma_\epsilon^2 = 2.5$	
	$\hat{\sigma^2}_{\epsilon OLS}$	$\hat{\sigma^2}_{\epsilon MME}$	$\hat{\sigma^2}_{\epsilon OLS}$	$\hat{\sigma^2}_{\epsilon MME}$	$\hat{\sigma^2}_{\epsilon OLS}$	$\hat{\sigma^2}_{\epsilon MME}$
$\sigma_u^2 = 1.5$	17.123633	19.105339	17.503852	19.194230	20.608246	20.832436
$\sigma_u^2 = 2.0$	24.857297	25.933181	23.392068	24.529332	27.502203	27.901681
$\sigma_u^2 = 2.5$	31.597691	33.130699	30.194104	32.498510	29.345670	31.759968
$\sigma_u^2 = 3.0$	32.975182	37.525819	32.044992	37.391702	35.729896	38.775078
$\sigma_u^2 = 4.5$	44.576956	56.048866	47.821178	57.611808	49.167870	56.316500
$\sigma_u^2 = 6.0$	48.857455	71.243179	46.908839	70.395988	65.226202	74.410970
$\sigma_u^2 = 8.0$	70.609430	98.734220	80.935690	98.421423	83.344002	92.865456
$\sigma_u^2 = 10.0$	61.902354	114.062432	85.163824	116.435787	68.716872	116.415571

and the difference when using $\hat{\sigma^2}_{\epsilon MME}$ is greater than that for $\hat{\sigma^2}_{\epsilon OLS}$.

5   Conclusion

We have presented some empirical results on the case of non-normality under the errors-in-variables model, and have shown how non-normality in the observation error affects the estimation of the regression parameters β and σ_ϵ^2. We have compared the errors in different metrics when using different distributions and different estimators $\hat{\beta}$ and $\hat{\sigma^2}_\epsilon$.

Future work can be extended upon this topic. We have only presented the empirical 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

7 Appendix

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

```python
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 = np.array([])
            sqerr_beta = np.array([])
            bias_beta_adj = np.array([])
            sqerr_beta_adj = np.array([])

            bias_sigma = np.array([])
            sqerr_sigma = np.array([])
            bias_sigma_adj_v2 = np.array([])
            sqerr_sigma_adj_v2 = 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_est = np.cov(x_obs, y)[0][1]/np.var(x_obs, ddof=1)
    beta_est_adj = np.cov(x_obs, y)[0][1]/np.var(x_obs, ddof=1) * np

    vx = np.var(x_obs, ddof=1) - var_u
    vu = var_u
    l = vx/(vx+vu);

    sigma_est = np.sum((y - beta_est * x_obs) ** 2) / (n-1)
    sigma_est_adj_v2 = np.sum((y - beta_est * x_obs) ** 2) / (n-1) -

    bias_beta = np.append(bias_beta, (beta_est-secret_beta))
    sqerr_beta = np.append(sqerr_beta, (beta_est-secret_beta)**2)
    bias_beta_adj = np.append(bias_beta_adj, (beta_est_adj-secret_be
    sqerr_beta_adj = np.append(sqerr_beta_adj, (beta_est_adj-secret_l

    bias_sigma = np.append(bias_sigma, sigma_est-var_eps)
    sqerr_sigma = np.append(sqerr_sigma, (sigma_est-var_eps)**2)
    bias_sigma_adj_v2 = np.append(bias_sigma_adj_v2, sigma_est_adj_v2
    sqerr_sigma_adj_v2 = np.append(sqerr_sigma_adj_v2, (sigma_est_ad

print("var_u=%.2f, var_eps=%.2f:   OLS: MBE = %f, MSE = %f; Sigma MBE

print("var_u=%.2f, var_eps=%.2f:  MME: MBE = %f, MSE = %f; Sigma MBE

beta_dict[(var_u, var_eps)] = {}
beta_dict[(var_u, var_eps)]['bias_ols'] = bias_beta.mean()
beta_dict[(var_u, var_eps)]['sqerr_ols'] = sqerr_beta.mean()
beta_dict[(var_u, var_eps)]['bias_ols_corr'] = bias_beta_adj.mean()
beta_dict[(var_u, var_eps)]['sqerr_ols_corr'] = sqerr_beta_adj.mean(

sigma_dict[(var_u, var_eps)] = {}
sigma_dict[(var_u, var_eps)]['bias_ols'] = bias_sigma.mean()
sigma_dict[(var_u, var_eps)]['sqerr_ols'] = sqerr_sigma.mean()
sigma_dict[(var_u, var_eps)]['bias_ols_corr_v2'] = bias_sigma_adj_v2
sigma_dict[(var_u, var_eps)]['sqerr_ols_corr_v2'] = sqerr_sigma_adj_v

return beta_dict, sigma_dict
```

References

[1] Steve Pischke. *Lecture Notes on Measurement Error*. 2007.

[2] Jonathan Gillard. *Method of Moments Estimation in Linear Regression with Errors in both Variables*. 2014.