Economic Growth Centre
Economics, School of Social Sciences Nanyang
Technological University
14 Nanyang Drive
Singapore 637332

COULD REGRESSION OF STATIONARY SERIES BE SPURIOUS?

Wing-Keung Wong
Yushan Cheng
Mu Yue

February 2025

EGC Report No: 2025/03

HSS-04-90A
Tel: +65 67906073
Email: D-EGC@ntu.edu.sg

Economic Growth Centre Working Paper Series

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE | Economic Growth Centre School of Social Sciences

Could regression of stationary series be spurious?

Wing-Keung Wong

Department of Finance, Fintech Center, and Big Data Research Center,

Asia University

Department of Medical Research, China Medical University Hospital, Taiwan

Department of Economics and Finance, The Hang Seng University of Hong Kong

Yushan Cheng

School of Mathematics and Statistics, Xi'an Jiaotong University, China

Mu Yue *

Engineering Systems and Design, Singapore University of Technology & Design

January 18, 2025

Acknowledgments: The first author would like to thank Robert B. Miller and Howard E. Thompson for their continuous guidance and encouragement. This research has been supported by Xi'an Jiaotong University, University of Canberra, Asia University, China Medical University Hospital, The Hang Seng University of Hong Kong, Research Grants Council (RGC) of Hong Kong (project number 12500915), and Ministry of Science and Technology (MOST, Project Numbers 106-2410-H-468-002 and 107-2410-H-468-002-MY3), Taiwan. This article was done when the first author visited the Nanyang Technological University in 2024.

Abstract:

Most of the literature on spurious regression has found that regression of independent and (nearly) non-stationary time series could result in spurious outcomes. Very few studies address

**Corresponding author: Mu Yue, Engineering Systems and Design, Singapore University of Technology, 8 Somapah Rd, Singapore 487372, email: yuemu.moon@gmail.com.

the issue of whether regression of stationary time series could also result in spurious outcomes and the study is not comprehensive and thorough. To bridge the gap in the literature, we first conjecture that regression of stationary time series could also result in spurious outcomes. We then examine whether the conjecture holds by providing a comprehensive and thorough study. We further provide a remedy algorithm to correct the spurious problem and improve the interpretability of the model. Extensive simulations are carried out to support our conjecture and demonstrate the effectiveness of the remedy. To demonstrate the applicability of our proposed approach and address the issue of spuriousness, we conduct a numerical analysis and demonstrate the usefulness of our proposed remedy algorithm.

Keywords: spurious regression, stationarity, non-stationarity, autoregressive model.

JEL Classification: C01, C15, C22, C58, C60

1 Introduction

Time series data refers to the collection of observations obtained through repeated measurements over time, usually at equally spaced time points. One important goal for time series analysis is modelling the data to provide useful insight about the underlying processes driving the series and forecast future outcomes of the series given the observed values. To this end, many useful models are proposed, such as the linear autoregressive moving average (ARMA) model, the autoregressive integrated moving average (ARIMA) model, the nonlinear threshold autoregressive (TAR) model (Tsay, 1989), the constant conditional correlation (CCC) model (Nakatani and Teräsvirta, 2009), and the conditional heteroscedasticity model (Tse and Tsui, 2002). There are also popular modelling methods based on exponential smoothing (Gardner Jr, 1985), including simple exponential smoothing (SES) and holt's winter exponential smoothing (HWES). We refer to Brockwell and Davis (2002) for an overview of time series modeling.

Many time series models are based on the assumption of stationarity. The statistical properties of a stationary data system do not change over time and the overall behavior of the system is consistent. In another words, different sections of the series will look roughly the same at intervals of the same length. Therefore it does not matter when we start to observe the series. In practice, the partial autocorrelation function (PACF) plot (Shao and Lund, 2004) and Augmented Dickey-Fuller (ADF) test (Mushtaq, 2011) can be used to test the stationarity assumption.

When the time series data are stationary, describing the relationship of two time series and making predictions of one process from the other is meaningful. Simple linear regression is one of the simplest predictive modelling algorithms which forecast the time series of interest assuming that it has a linear relationship with other time series. It is widely used in practice and adapts naturally to complex forecasting tasks. For example, the static Phillips curve which regresses the annual inflation rate and the unemployment rate are used to study the contemporaneous trade-off between them. Recently many authors proposed machine learning algorithms such as XG Boost (Xia et al., 2020), long short term memory (LSTM) (Malhotra et al., 2015), recurrent neural networks (Hewamalage et al., 2021), among others. In this post, we focus on the regression approach and our discussion can be further extended to more sophisticated procedures.

We investigate the spurious regression problem in this paper. Particularly, if a pair of time series are not causally related but appear to be moderately or even strongly correlated, due to either coincidence or the presence of a certain hidden factor, then the inference of the regression analysis could be unreliable. Key features of a spurious linear regression include spuriously significant regression coefficients, large R-squares (R^2), and low Durbin–Watson ratios (Phillips, 2009). Granger and Newbold (1974), Engle and Granger (1987), Granger (1981), Phillips (1986), Entorf (1997), Pesaran et al. (1999) and Westerlund (2008) all pointed out that regression of independent non-stationary time series could be spurious. Recently, Tu and Wang (2022) investigates the effect that spurious functional-coefficient regression involving nonstationary regressors has on model diagnostics. See, for example, Ventosa-Santaulária (2009), Phillips (2009) and Chen and Tu (2019) for a complete overview of spurious regression.

Most earlier works on spurious regression found that regression of independent and (nearly) non-stationary series could result in spurious outcomes. Very few studies address the issue of whether regression of stationary series could also result in spurious outcomes and the theoretical and applied studies are not comprehensive and thorough. To bridge the gap in the literature, we first conjecture that regression of stationary time series could also result in spurious outcomes.

We then examine this conjecture by extending the work by Granger (2001), Granger IV et al. (2001), Stock and Watson (2007), Hill et al. (2008), Wooldridge (2006), McCallum (2010), among others. In particular, we study the relationship between two stationary time series by performing a simple linear regression and subsequently testing the slope estimates. Thereafter, we conduct numerical investigation of the spurious regression. In this paper, we use AR(1) time series because the structure is simple and useful in a wide range of contexts. Extensive simulation studies are carried out and we show that under some situations, regression of two independent and stationary series could have a spurious problem. To resolve the issue, we further propose a remedy algorithm and demonstrate its applicability through further numerical analysis.

Granger and Newbold (1974), Ferson et al. (2003, 2008), Deng (2014) and others extended the earlier work by Yule (1926) and showed that it could be a spurious regression even though the variables are not autocorrelated. Agiakloglou (2013) studies the spurious regression problem for two independent stationary and non-stationary processes and Kim et al. (2004) examines the problem of spurious regression for two independent variables in which one has a linear trend and another one is $I(0)$. On the other hand, Tsay and Chung (2000) extends the theory of the spurious regression to including the long memory fractionally integrated processes and Abeysinghe (1994) extends the theory by considering seasonal dummies in regressions and when seasonality dummies follow seasonal unit roots.

The asymptotic theory of the estimation and testing for spurious regression has been well developed. For example, Phillips (1986), Sun (2004), Agiakloglou (2013) and others all provided asymptotic results for spurious regressions. Moreover, Marmol (1995, 1998) use fractional differencing techniques to develop statistics for spurious regressions of $I(d)$ processes and Cappuccio and Lubian (1997) establishes some properties for the asymptotic distribution theory of spurious regression with $I(1)$ processes and error term that are long-memory stationary errors. Kao (1999) develops some asymptotic properties for the estimators and statistics for the spurious regressions in panel data. In addition, Giles (2007) develops some properties for some tests for both homoskedasticity and normality of the innovation in a spurious regression, Choi et al. (2008) uses the generalized least squares approach to develop some asymptotic properties for spurious regressions, and Deng (2014) develops a diagnostic test to test whether there is any spurious regression bias for any empirical study. Ventosa-Santaulária (2009) reviews the theory of the spurious regression. There are also plenty of empirical studies on spurious regression. For example, see Ferson et al. (2003), Choi et al. (2008), Chatelain and Ralf (2014) and Deng (2014).

2 Spurious linear regression model

In this paper, we impose the following simple linear regression model for two series $\{Y_t, X_t, t = 1, \cdots, N\}$:

$$Y_t = \alpha + \beta X_t + u_t, \tag{2.1}$$

where u_t is an error term, N is the total length of the series, α is the intercept parameter, and β is the slope parameter. We assume that u_t has mean 0 and variance σ_u^2. The zero mean condition is easily satisfied after we include the intercept α, while the constant variance condition can be achieved if Y_t and X_t have constant variances (eg. being stationary processes) and $cov(Y_t, X_t)$ does not depend on time.

The simple linear regression model assumed in (2.1) depicts the relationship between two variables Y_t and X_t, and one important hypothesis for the simple linear regression is

$$H_0 \; : \; \beta = 0 \quad \text{versus} \quad H_1 \; : \; \beta \neq 0 \; . \tag{2.2}$$

If the null hypothesis H_0 in (2.2) is true, then Y_t does not change with X_t. In this case, we claim that Y_t does not depend on the value of X_t, implying that X_t and Y_t have no linear association. If the alternative hypothesis H_1 is true, then change in X_t is associated with change in Y_t linearly.

To test whether the null hypothesis H_0 in (2.2) is true, the following T test is commonly used:

$$T = \frac{\hat{\beta} - \beta}{SE(\hat{\beta})} = \frac{\hat{\beta}}{SE(\hat{\beta})} \; , \tag{2.3}$$

where

$$\hat{\beta} = \frac{\sum_{t=1}^{N} (X_t - \bar{X})(Y_t - \bar{Y})}{\sum_{t=1}^{N} (X_t - \bar{X})^2} \tag{2.4}$$

is the least squares estimate (LSE) of β in which $\bar{X} = \sum_{t=1}^{N} X_t/N$, $\bar{Y} = \sum_{t=1}^{N} Y_t/N$, and SE($\hat{\beta}$) is the standard error of the LSE calculated by:

$$SE(\hat{\beta}) = \sqrt{\frac{\sum_{t=1}^{N} (\hat{u}_t^2)/(N-2)}{\sum_{t=1}^{N} (X_t - \bar{X})^2}} \; , \tag{2.5}$$

with $\hat{u}_t = (Y_t - \hat{\alpha} - \hat{\beta} X_t)^2$. It is well known that the test statistic T follows a t-distribution with $N-2$ degrees of freedom if the null hypothesis H_0 is true and normality is assumed.

The *goodness-of-fit* or R^2 of the fitted model is

$$R^2 = \frac{SSR}{SSTO} = 1 - \frac{SSE}{SSTO} \ , \tag{2.6}$$

where $SSR = \sum_{i=1}^{N} (\hat{Y}_i - \bar{Y}_i)^2$, $SSTO = \sum_{i=1}^{N} (Y_i - \bar{Y}_i)^2$ and $SSE = \sum_{i=1}^{N} (Y_i - \hat{Y}_i)^2$. The R^2 is a proportion between 0 and 1. The model with a larger R^2 value will have more variation in Y_t that can be explained, and therefore, it is a better fitted model.

A spurious relationship means that a detected empirical relationship actually does not make any sense. Granger and Newbold (1974) noted that if

$$Y_t \sim I(1) \quad \text{and} \quad X_t \sim I(1) \ , \tag{2.7}$$

then regression in (2.1) could be spurious. In another word, when Y_t and X_t are random walks and non-stationary, even if they are completely independent, the test statistic T in (2.3) from ordinary least square regression could be significant and misleadingly indicate a good fit of linear model.

In this paper, we complement the work by Granger and Newbold (1974) by studying the following conjecture:

Conjecture 1 *Even if both Y_t and X_t are stationary, regression in (2.1) could still be spurious for an empirical study.*

If Conjecture 1 holds, the following question arises: Is the conventional T test in ordinary least square regression a valid method to examine the relationship between two stationary time series? Our response will depend on a series of numerical analysis.

3 Simulation

3.1 Simulation Model Setup

We now consider the simple linear regression in (2.1) between two unrelated stationary AR(1) series X_t and Y_t in which

$$X_t = \alpha_1 X_{t-1} + \varepsilon_t, \quad \text{and} \quad Y_t = \alpha_2 Y_{t-1} + e_t, \tag{3.1}$$

in which $|\alpha_1|$ and $|\alpha_2|$ are strictly less than 1 and, for simplicity, we assume both ε_t and e_t are independent innovations with the following continuous probability density function

$$f(a;p) \propto \frac{1}{\sigma}\left\{1+\frac{a^2}{k\sigma^2}\right\}^{-p} \qquad (-\infty < a < \infty)\ , \tag{3.2}$$

where $k = 2p-1$ and $p \geq 2$. We note that $E(a) = 0$, $Var(a) = \sigma^2$, and $t = \sqrt{(k/\nu)}\,(a/\sigma)$ has Student's t distribution with $\nu = 2p-1$ degrees of freedom. For $1 \leq p < 2$, k is equal to 1 and in this case, σ in (3.2) is simply a scale parameter. When $\nu \to \infty$, $f(a)$ becomes a normal distribution. The AR(1) model specifies that the output variable depends linearly on its own preceding value where *innovation* adds new information to the series but is uncorrelated with past values of time series.

Without loss of generality, we will consider three different factors that could affect the behavior of the two time series. Firstly, we consider the distribution of the error terms which is a main factor affecting the characteristics of time series. In particular, we consider both ε_t and e_t follow $N(0,1)$, $t(5)$, $t(2)$ and $t(1)$. Secondly, we vary the lengths of times series since longer series will include more information than shorter one. In this paper, we simulate the time series with $N = 100$, 200, 400 and 800. Finally, we consider the different values of auto-regressive parameters α_1 and α_2. In our experiment, we consider two sets of parameter values $A^- = \{-0.1, -0.3, -0.5, -0.7, -0.9\}$ and $A^+ = \{0.1, 0.3, 0.5, 0.7, 0.9\}$. We consider the values of $(\alpha_1, \alpha_2) \in (A^+, A^+)$, (A^-, A^-), (A^+, A^-) and (A^-, A^+). With 4 different error distributions, 4 different time series lengths, and $5 \times 5 \times 4 = 100$ combinations of α_1 and α_2 values, there are in total $4 \times 4 \times 100 = 1600$ cases.

As X_t and Y_t are generated independently, they are completely unrelated. Thus, one may expect the slope coefficient estimator $\hat{\beta}$ is not significantly different from zero.

For each case (different error distributions, different time series lengths, different combinations of α_1 and α_2) described in section 3.1, we consider the following computational procedure:

Procedure 1

---

a. Set the number of simulations to be 10000.

b. Fit model (2.1) for each simulation. Obtain 10000 $\hat{\beta}$'s and their corresponding p-values.

c. Plot the empirical distribution of $\hat{\beta}$'s. The standard errors and t-statistics of $\hat{\beta}$'s are also closely monitored in case of any abnormality.

d. Use the T test defined in (2.3) to test whether the null hypothesis H_0 in (2.2) is true. H_0 is rejected if p-value for the T test is less than 0.05. Calculate the proportion of significant $\hat{\beta}$'s or proportion of p-values that are less than 0.05 among the 10000 fitted linear regression models in each case.

The above procedure helps us to examine whether the T statistic in (2.3) for the model in (2.1) indeed follows a Student's t-distribution. When X_t and Y_t are unrelated, the null hypothesis that β coefficient is zero should be rejected less than 5% of the total simulations at the significance level of 0.05. If the T test is suitable, that is, $\hat{\beta}$'s follow the Student's t-distribution, the rejection rate will be exactly 0.05.

3.2 Simulation Results

Detailed results are provided in the Appendix A - E. Under each error distribution specified in section 3.1, we find that the histograms of $\hat{\beta}$ are similar. Thus we only present the histogram of $\hat{\beta}$'s for $\alpha_1 \in A^+$ and $\alpha_2 = 0.7$ when both ε_t and $e_t \sim$ N(0,1) (Appendix A). The probability of rejecting the null hypothesis for all the cases are recorded in Appendices B - E.

We note that the simulation results of $(\alpha_1, \alpha_2) \in (A^-, A^-)$ are similar to those of $(\alpha_1, \alpha_2) \in (A^+, A^+)$, while the results of $(\alpha_1, \alpha_2) \in (A^+, A^-)$ are similar to those of $(\alpha_1, \alpha_2) \in (A^-, A^+)$. Thus, we choose $(\alpha_1, \alpha_2) \in (A^+, A^+)$ and $(\alpha_1, \alpha_2) \in (A^+, A^-)$ as representatives to report in the following.

Case I. $(\alpha_1, \alpha_2) \in (A^+, A^+)$

When $(\alpha_1, \alpha_2) \in (A^+, A^+)$, the empirical distributions of $\hat{\beta}$'s are shown in Figures A1. In general, the histogram is very close to a bell shape. The overall shape of the distribution does not differ for different values of α_1 and α_2 and different lengths of time series. This suggests that a t-test could be valid for testing the significance of $\hat{\beta}$ since the t-test usually requires the sampling distribution to be close to a normal distribution.

When both ε_t and $e_t \sim$ N(0,1), t(5) or t(2), the average rejection rate is increasing with the increase of α_1 and α_2. In particular, when α_1 and α_2 increase from 0.1 to 0.9, the average rejection rate increases from 5% to 50%. When both ε_t and $e_t \sim$ t(1), the average rejection rate is also increasing with the increase of α_1 and α_2, ranging from 0.0035 to 0.407. The results thus support Conjecture 1 that regression between two stationary time series could be spurious. In addition, the rejection rate is effected by the value of α_1 and α_2. With the increasing of α_1 and α_2, the rejection rate increases from about 3.5% to about 50%. When both α_1 and α_2 approach 1 (hence both processes are close to random walks), the regression between two stationary time series is more likely to become spurious. Moreover, when α_1 and α_2 are equal, the effect from the lengths of time series is very small. Only when both ε_t and $e_t \sim$ t(1) and when both α_1 and α_2, are close to zero, the rejection rates of smaller α_1 and α_2 may be less than 0.05.

Case II. $(\alpha_1, \alpha_2) \in (A^+, A^-)$

When $(\alpha_1, \alpha_2) \in (A^+, A^-)$, the rejection rate are shown in Appendix D (Tables D1 - D4). When both ε_t and $e_t \sim$ N(0,1), t(5), t(2) or t(1), almost all rejection rates are smaller than the 5% level of significance. The rejection rates also seem to be decreasing with magnitudes of α_1 and α_2 increasing. Thus, when Y_t and X_t follow stationary AR(1) models with opposite signs on the AR(1) parameter, there seems to be no spurious issue, suggesting that Conjecture 1 is not supported in this case. However, the t-test is still inappropriate since the empirical rejection rate could be much lower than the nominal level.

From the above simulation, we can conclude that when Y_t and X_t follow stationary AR(1) models with the same signs on the AR(1) parameters α_1 and α_2, the regression could be spurious. In all cases, the traditional t-test is invalid and the test results are questionable.

Remark. We briefly discuss the distribution issue of the estimated spurious coefficient $\hat{\beta}$ in equation (2.1). For a fixed-design study, the regressors are usually assumed to be non-random, and thus, the distribution of $\hat{\beta}$ can be easily derived to be the normal distribution. Hence, the test for regression coefficient follows the well-known t-distribution. However, in the study of two seemingly related stochastic processes, both dependent and independent variables should be treated as random. It is thus inappropriate to claim that $\hat{\beta}$ follows normal distributed. The analytic distribution involves a functional of Wiener process as shown in Phillips (1986). In general, the asymptotic distribution of $\hat{\beta}$ cannot be written in a simple form and can only be simulated using numerical methods. However, the simulated distribution could also involve non-trivial numerical integration and may not be very accurate for a practical sample. We do not strongly recommend to use such a distribution for the subsequent hypothesis tests and the corresponding inference.

4 Remedy

Simulation studies in section 3 have shown that Conjecture 1 could hold for many practical settings. In this section, we will investigate whether there is any remedy to solve this problem. We propose the following conjecture:

Conjecture 2 *The spurious problem of regressing a stationary series to another independent stationary series could be overcome after we apply a transformation to the original series.*

Because both X_t and Y_t are AR(1) time series, autocorrelation might appear in the residuals of model (2.1); that is:

$$\text{corr}(u_t, u_{t-1}) = \rho \neq 0 \ , \tag{4.1}$$

in which ρ represents the lag one autocorrelation within the residuals of regression model (2.1). We can detect the presence of autocorrelation in the error by using the Durbin-Watson test; that is, to check whether $\rho = 0$. The Durbin-Watson test is very useful in many applications since the first-order autocorrelation is very likely to be present in any time series data where the preceding value tends to have an impact on the immediate following value. A very low Durbin-Watson test statistic could be the signal of positive autocorrelation while a very large test statistic indicates a possible negative autocorrrelation. Autocorrelation in the residuals need to be removed in order to decide if X_t and Y_t are truly linearly related.

To circumvent the problem of the autocorrelation in the error term, we modify the algorithm introduced by Tiku et al. (1999) and Wong and Bian (2005) as follows:

Procedure 2

---

a. Perform the Durbin-Watson test on the residuals for each of the linear regression performed in Procedure 1.

b. If H_0: $\rho = 0$ is rejected, estimate the autocorrelation by the sample moment estimator $\hat{\rho}$

and transform the original X_t and Y_t to X_t' and Y_t' through the following equations:

$$\begin{aligned} Y_t' &= (1-\hat{\rho}B)Y_t = Y_t - \hat{\rho}Y_{t-1} \ , \\ X_t' &= (1-\hat{\rho}B)X_t = X_t - \hat{\rho}X_{t-1} \ , \end{aligned}$$

where B is the backshift operator. If there is not enough evidence to reject H_0, keep the original rejection rate and exit the loop.

c. Fit the following linear model between Y_t' and X_t':

$$Y_t' = \gamma_0 + \gamma_1 X_t' + \epsilon_t^* \ , \tag{4.2}$$

which is defined similarly as model (2.1). **d.** Perform the Durbin-Watson test on the residuals of the fitted Model (4.2). If H_0 is rejected, go back to Step b to remove the autocorrelation in Y_t' and X_t' in the same way, and continue the loop. If not, record the new rejection rate and exit the loop. If after 3 loops, H_0 is stilled rejected, this means that there is still autocorrelation in the residuals after transforming 3 times, then exit the loop and keep the rejection rate of the last iteration.

We re-do the analysis for the simulated data using Procedure 2 and record the rejection rates before and after transformation in Appendix F (Tables F1 - F4). The results show that the transformation algorithm is effective and can significantly reduce the probability of rejecting the null hypothesis. The empirical rejection rates are now all close to the nominal 0.05 level and the results are not sensitive to the length of the sequence or the distribution of errors.

5 Real data analysis

Financial time series analysis plays an important role in hedging market risks and optimizing investment decisions. In this section, we analyze the Alphabet Inc. closing stock price (GOOG) and Edison International closing stock price (EIX) from 7-Feb-2005 to 7-July-2005. Alphabet Inc. is an American multinational technology conglomerate holding company headquartered in Mountain View, California and also the parent company of Google. Edison International is a public utility holding company based in Rosemead, California. The data can be downloaded from `https://sg.finance.yahoo.com/quote/YHOO/history/`.

The time series for GOOG and EIX prices are plotted in the 1st row of Figure 1. Consecutive values appear to follow one another fairly closely, suggesting an autoregression model could be appropriate to model both time series. We next look at the PACF plots for the two stock price data which are shown in the 2nd row of Figure 1. Both PACF plots show a significant spike at a lag of 1 and much lower spikes for the subsequent lags. Thus, an AR(1) model would likely be suitable for the data. ADF test is further conducted to show that the AR(1) process is stationary. The first order lag plots shown in the 3rd row of figure 1 validate the feasibility for AR(1) model for both data.

Now, we regress the GOOG price (Y_t) on the EIX price (X_t) with a linear regression model (2.1). The p-value of the coefficient for EIX price is smaller than 2e-16 and the t-test is thus highly significant. To explore whether the regression is spurious, we apply the remedy algorithm in section 4. The p-value of the coefficient on the transformed time series model increases to 0.16 after one loop, suggesting that the original regression might be spurious and the significant output may be misleading. The scatter plots for GOOG price and EIX price with the fitted regression line before and after remedy are shown in the last row of Figure 1.

6 Conclusion

This research note aims to find out whether the regression of two stationary time series is spurious and what remedy can be applied to solve the spuriousness issue. To this end, we set up a conjecture that even if both dependent and independent variables are stationary, regression of the variables could also be spurious. To test whether the conjecture holds, we simulated independent AR(1) processes and checked the proportion of rejected H_0: $\beta = 0$ in model (2.1). Results in section 3 successfully demonstrate that the conjecture could hold for a wide range of settings and the conventional t-test in the OLS linear regression is not helpful in detecting spurious regression. Our findings imply that when one runs a regression between time series with unknown structures or properties, the probability of incorrectly concluding a significant association could be substantial when using the traditional t-test. We suggest to the practitioners that they should check carefully to avoid spurious regression in their analysis. One way of detecting the spurious problem is to apply the transformed model fitting procedure proposed in section 4.

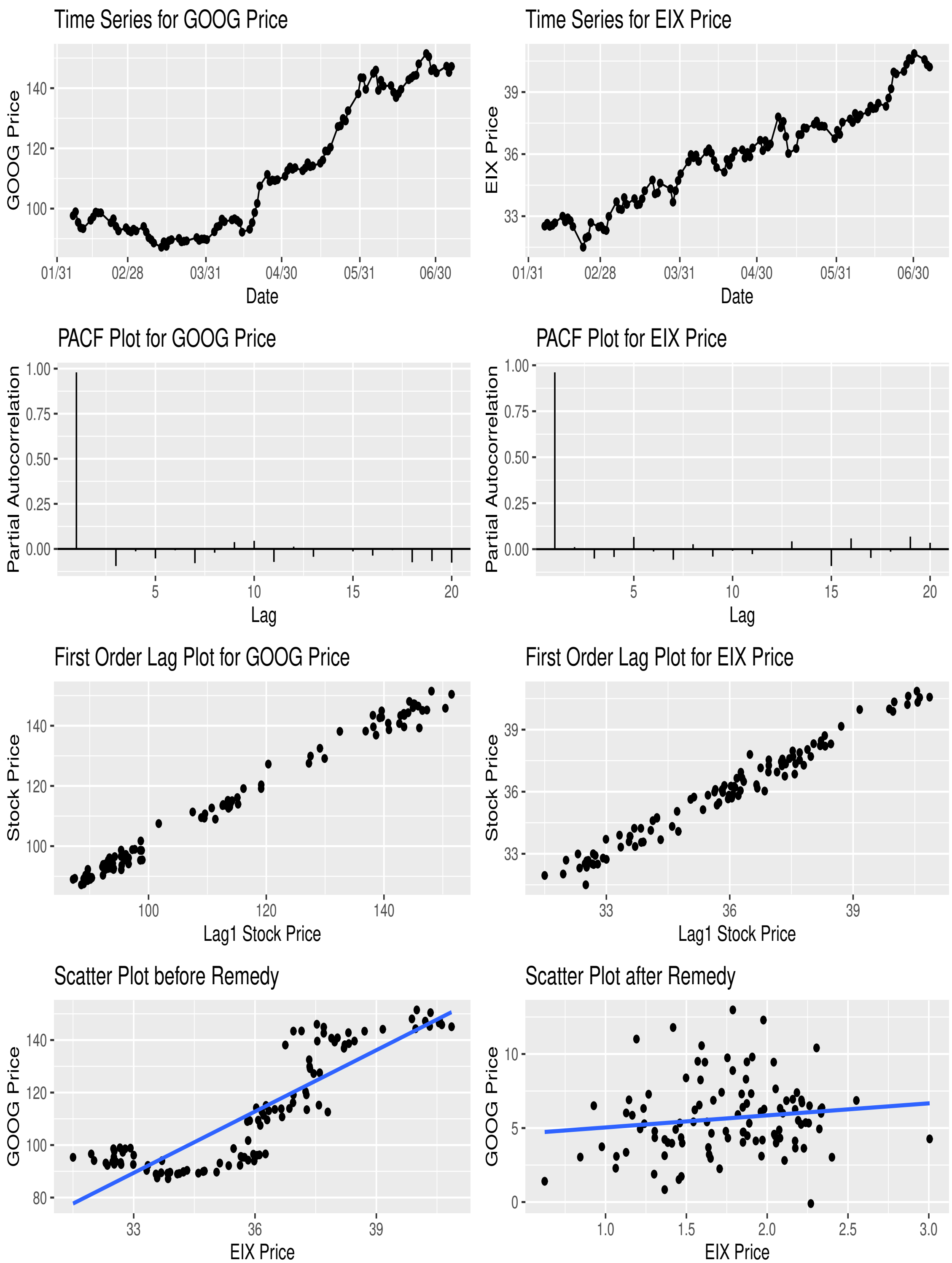

Figure 1: The 1st row shows the time series for GOOG price and EIX price from 7-Feb-2005 to 7-July-2005; The 2nd row shows the PACF plots for GOOG price and EIX price; The 3rd row shows the first order lag plot for GOOG price and EIX price; The last row shows the scatter plots for GOOG price and EIX price with the fitted regression line before and after remedy.

Extensions of our work include improving the approach proposed in our paper to a broader class of time series models. Future research could also include adding more independent variables in (2.1) to analyze whether the spurious outcomes could present in multivariate linear regression, investigating whether a combination of stationary and non-stationary independent variables could lead to similar outcomes, and examining whether there is any other appropriate test for higher-order models that could perform as well as in this context for the simple linear regression. Extensions could also include discovering other ways of transforming models to achieve better results than what is proposed in our paper.

References

Abeysinghe, T. (1994), 'Deterministic seasonal models and spurious regressions', *Journal of Econometrics* **61**(2), 259–272.

Agiakloglou, C. (2013), 'Resolving spurious regressions and serially correlated errors', *Empirical Economics* **45**(3), 1361–1366.

Brockwell, P. J. and Davis, R. A. (2002), *Introduction to time series and forecasting*, Springer.

Cappuccio, N. and Lubian, D. (1997), 'Spurious regressions between i (1) processes with long memory errors', *Journal of Time Series Analysis* **18**(4), 341–354.

Chatelain, J.-B. and Ralf, K. (2014), 'Spurious regressions and near-multicollinearity, with an application to aid, policies and growth', *Journal of macroeconomics* **39**, 85–96.

Chen, Y. and Tu, Y. (2019), 'Is stock price correlated with oil price? spurious regressions with moderately explosive processes', *Oxford Bulletin of Economics and Statistics* **81**(5), 1012–1044.

Choi, C.-Y., Hu, L. and Ogaki, M. (2008), 'Robust estimation for structural spurious regressions and a hausman-type cointegration test', *Journal of Econometrics* **142**(1), 327–351.

Deng, A. (2014), 'Understanding spurious regression in financial economics', *Journal of Financial Econometrics* **12**(1), 122–150.

Engle, R. F. and Granger, C. W. (1987), 'Co-integration and error correction: representation, estimation, and testing', *Econometrica: journal of the Econometric Society* pp. 251–276.

Entorf, H. (1997), 'Random walks with drifts: Nonsense regression and spurious fixed-effect estimation', *Journal of Econometrics* **80**(2), 287–296.

Ferson, W. E., Sarkissian, S. and Simin, T. (2008), 'Asset pricing models with conditional betas and alphas: The effects of data snooping and spurious regression', *Journal of Financial and Quantitative Analysis* **43**(2), 331–353.

Ferson, W. E., Sarkissian, S. and Simin, T. T. (2003), 'Spurious regressions in financial economics?', *The Journal of Finance* **58**(4), 1393–1413.

Gardner Jr, E. S. (1985), 'Exponential smoothing: The state of the art', *Journal of forecasting* **4**(1), 1–28.

Giles, D. E. (2007), 'Spurious regressions with time-series data: further asymptotic results', *Communications in statistics—Theory and methods* **36**(5), 967–979.

Granger, C. W. (1981), 'Some properties of time series data and their use in econometric model specification', *Journal of econometrics* **16**(1), 121–130.

Granger, C. W. (2001), 'Spurious regressions in econometrics', *A Companion to Theoretical Econometrics, Blackwell Publishers* pp. 557–561.

Granger, C. W. and Newbold, P. (1974), 'Spurious regressions in econometrics', *Journal of econometrics* **2**(2), 111–120.

Granger IV, C. W., Hyung, N. and Jeon, Y. (2001), 'Spurious regressions with stationary series', *Applied Economics* **33**(7), 899–904.

Hewamalage, H., Bergmeir, C. and Bandara, K. (2021), 'Recurrent neural networks for time series forecasting: Current status and future directions', *International Journal of Forecasting* **37**(1), 388–427.

Hill, R. C., Griffiths, E. and Lim, C. (2008), 'Principles of econometrics, john miley & sons', *Inc., USA* pp. 695–697.

Kao, C. (1999), 'Spurious regression and residual-based tests for cointegration in panel data', *Journal of econometrics* **90**(1), 1–44.

Kim, T.-H., Lee, Y.-S. and Newbold, P. (2004), 'Spurious regressions with stationary processes around linear trends', *Economics Letters* **83**(2), 257–262.

Malhotra, P., Vig, L., Shroff, G., Agarwal, P. et al. (2015), Long short term memory networks for anomaly detection in time series, *in* 'Proceedings', Vol. 89, pp. 89–94.

Marmol, F. (1995), 'Spurious regressions between i (d) processes', *Journal of Time Series Analysis* **16**(3), 313–321.

Marmol, F. (1998), 'Spurious regression theory with nonstationary fractionally integrated processes', *Journal of Econometrics* **84**(2), 233–250.

McCallum, B. T. (2010), 'Is the spurious regression problem spurious?', *Economics Letters* **107**(3), 321–323.

Mushtaq, R. (2011), 'Augmented dickey fuller test', *http://dx.doi.org/10.2139/ssrn.1911068* .

Nakatani, T. and Teräsvirta, T. (2009), 'Testing for volatility interactions in the constant conditional correlation garch model', *The Econometrics Journal* **12**(1), 147–163.

Pesaran, M. H., Shin, Y. and Smith, R. P. (1999), 'Pooled mean group estimation of dynamic heterogeneous panels', *Journal of the American statistical Association* **94**(446), 621–634.

Phillips, P. C. (1986), 'Understanding spurious regressions in econometrics', *Journal of econometrics* **33**(3), 311–340.

Phillips, P. C. (2009), 'Local limit theory and spurious nonparametric regression', *Econometric Theory* **25**(6), 1466–1497.

Shao, Q. and Lund, R. (2004), 'Computation and characterization of autocorrelations and partial autocorrelations in periodic arma models', *Journal of Time Series Analysis* **25**(3), 359–372.

Stock, J. H. and Watson, M. W. (2007), 'Why has us inflation become harder to forecast?', *Journal of Money, Credit and banking* **39**, 3–33.

Sun, Y. (2004), ‘A convergent t-statistic in spurious regressions’, *Econometric Theory* **20**(5), 943–962.

Tiku, M. L., Wong, W.-K. and Bian, G. (1999), ‘Estimating parameters in autoregressive models in non-normal situations: Symmetric innovations’, *Communications in Statistics-Theory and Methods* **28**(2), 315–341.

Tsay, R. S. (1989), ‘Testing and modeling threshold autoregressive processes’, *Journal of the American statistical association* **84**(405), 231–240.

Tsay, W.-J. and Chung, C.-F. (2000), ‘The spurious regression of fractionally integrated processes’, *Journal of Econometrics* **96**(1), 155–182.

Tse, Y. K. and Tsui, A. K. C. (2002), ‘A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations’, *Journal of Business & Economic Statistics* **20**(3), 351–362.

Tu, Y. and Wang, Y. (2022), ‘Spurious functional-coefficient regression models and robust inference with marginal integration’, *Journal of Econometrics* **229**(2), 396–421.

Ventosa-Santaulària, D. (2009), ‘Spurious regression’, *Journal of Probability and Statistics* **2009**.

Westerlund, J. (2008), ‘Panel cointegration tests of the fisher effect’, *Journal of Applied Econometrics* **23**(2), 193–233.

Wong, W.-K. and Bian, G. (2005), ‘Estimating parameters in autoregressive models with asymmetric innovations’, *Statistics & probability letters* **71**(1), 61–70.

Wooldridge, J. M. (2006), ‘Introduction to econometrics: A modern approach’, *Michigan State University. USA* .

Xia, Z., Xue, S., Wu, L., Sun, J., Chen, Y. and Zhang, R. (2020), ‘Forexgboost: passenger car sales prediction based on xgboost’, *Distributed and Parallel Databases* **38**(3), 713–738.

Yule, G. U. (1926), ‘Why do we sometimes get nonsense-correlations between time-series?–a study in sampling and the nature of time-series’, *Journal of the royal statistical society* **89**(1), 1–63.

Appendices

A The histogram of $\hat{\beta}$ for $\alpha_1 \in A^+, \alpha_2 = 0.7$, error $\sim$ N(0,1)

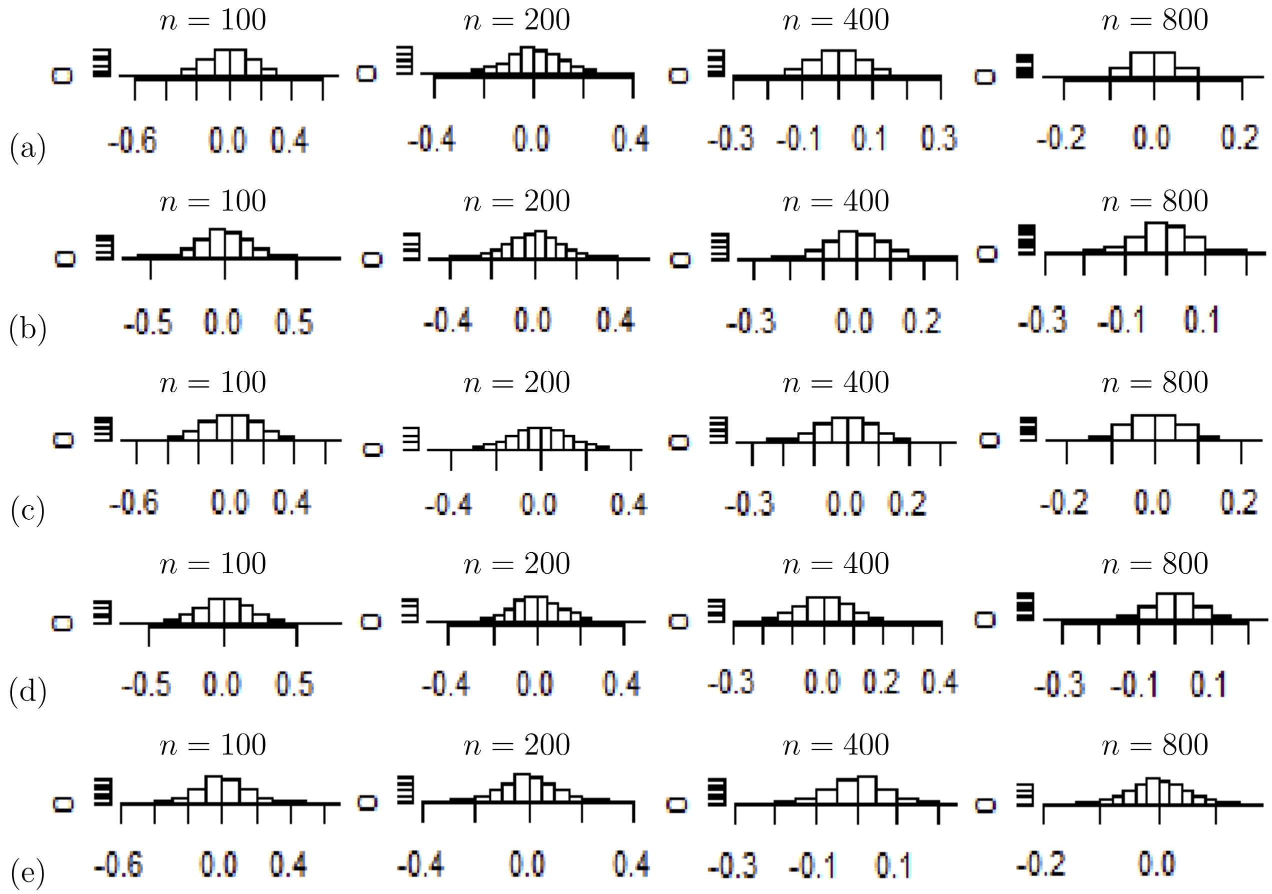

Figure A1: The histogram of $\hat{\beta}$ when errors $\sim$ N(0,1). Cases (a)-(e) represent results for $\alpha_1 = 0.1, 0.3, 0.5, 0.7, 0.9$ respectively.

B The rejection rate for $(\alpha_1, \alpha_2) \in (A^+, A^+)$

Table B1: Rejection Rate, error $\sim$ N(0,1)

α_2	α_1	N=100	N=200	N=400	N=800	Raverage	Saverage
	0.1	0.056	0.052	0.052	0.055	0.054	
	0.3	0.059	0.056	0.060	0.054	0.057	
0.1	0.5	0.063	0.063	0.061	0.066	0.063	0.063
	0.7	0.069	0.068	0.067	0.068	0.068	
	0.9	0.073	0.075	0.074	0.071	0.073	
Caverage		0.064	0.063	0.063	0.063	0.063	0.063
	0.1	0.060	0.057	0.056	0.061	0.058	
	0.3	0.071	0.074	0.074	0.075	0.073	
0.3	0.5	0.089	0.090	0.094	0.087	0.090	0.094
	0.7	0.107	0.113	0.112	0.115	0.112	
	0.9	0.136	0.129	0.138	0.136	0.135	
Caverage		0.092	0.093	0.095	0.095	0.094	0.094
	0.1	0.064	0.063	0.064	0.062	0.063	
	0.3	0.091	0.089	0.096	0.091	0.092	
0.5	0.5	0.128	0.129	0.129	0.131	0.129	0.136
	0.7	0.169	0.174	0.172	0.177	0.173	
	0.9	0.220	0.229	0.220	0.226	0.224	
Caverage		0.135	0.137	0.136	0.137	0.136	0.136
	0.1	0.064	0.068	0.067	0.069	0.067	
	0.3	0.112	0.110	0.115	0.112	0.112	
0.7	0.5	0.176	0.174	0.173	0.174	0.174	0.190
	0.7	0.240	0.250	0.249	0.253	0.248	
	0.9	0.340	?0.349	0.347	0.357	0.348	
Caverage		0.186	0.190	0.190	0.193	0.190	0.190
	0.1	0.070	0.070	0.067	0.074	0.070	
	0.3	0.133	0.137	0.137	0.133	0.135	
0.9	0.5	0.223	0.228	0.230	0.230	0.228	0.258
	0.7	0.328	0.342	0.340	0.346	0.339	
	0.9	0.501	0.519	0.522	0.531	0.518	
Caverage		0.251	0.259	0.259	0.263	0.258	0.258
Overall average		0.146	0.148	0.149	0.150	0.148	0.148

Note: ‘Raverage’ (stands for row average) is the average rejection rate for the same value of α_1 and α_2 but across different values of N ; ‘Caverage’ (stands for column average) is the average rejection rate for the same value of N and α_2 but across different values of α_1; ‘Saverage’ (stands for Situation average) is the average rejection rate for the same value of α_2 but across different values of α_1 and N; ‘Overall average’ in columns 3 - 6 is the overall average for each N for different values of α_1 and α_2 while ‘Overall average’ in the last two column is the overall average for all the cases in the entire table.

Table B2: Rejection Rate, error ∼ t(5)

α_2	α_1	N=100	N=200	N=400	N=800	Raverage	Saverage
	0.1	0.056	0.054	0.053	0.051	0.053	
	0.3	0.060	0.057	0.057	0.059	0.058	
0.1	0.5	0.068	0.061	0.068	0.059	0.064	0.063
	0.7	0.067	0.068	0.067	0.067	0.067	
	0.9	0.068	0.072	0.077	0.075	0.073	
Caverage		0.064	0.063	0.064	0.062	0.063	0.063
	0.1	0.059	0.061	0.055	0.060	0.059	
	0.3	0.076	0.076	0.074	0.065	0.073	
0.3	0.5	0.089	0.090	0.095	0.093	0.092	0.095
	0.7	0.112	0.116	0.113	0.115	0.114	
	0.9	0.138	0.139	0.137	0.134	0.137	
Caverage		0.095	0.096	0.095	0.093	0.095	0.095
	0.1	0.067	0.060	0.064	0.063	0.063	
	0.3	0.090	0.096	0.088	0.091	0.091	
0.5	0.5	0.124	0.129	0.128	0.127	0.127	0.135
	0.7	0.171	0.173	0.169	0.178	0.173	
	0.9	0.213	0.223	0.224	0.231	0.223	
Caverage		0.133	0.136	0.134	0.138	0.135	0.135
	0.1	0.067	0.067	0.069	0.065	0.067	
	0.3	0.112	0.114	0.112	0.111	0.112	
0.7	0.5	0.176	0.166	0.175	0.18	0.174	0.189
	0.7	0.249	0.243	0.249	0.254	0.249	
	0.9	0.334	0.349	0.345	0.343	0.343	
Caverage		0.188	0.188	0.190	0.190	0.189	0.189
	0.1	0.073	0.072	0.078	0.074	0.074	
	0.3	0.138	0.138	0.140	0.142	0.139	
0.9	0.5	0.225	0.228	0.216	0.229	0.224	0.260
	0.7	0.345	0.338	0.354	0.349	0.346	
	0.9	0.502	0.517	0.517	0.527	0.516	
Caverage		0.257	0.259	0.261	0.264	0.260	0.260
Overall average		0.147	0.148	0.149	0.150	0.148	0.148

Note: 'Raverage' (stands for row average) is the average rejection rate for the same value of α_1 and α_2 but across different values of N ; 'Caverage' (stands for column average) is the average rejection rate for the same value of N and α_2 but across different values of α_1; 'Saverage' (stands for Situation average) is the average rejection rate for the same value of α_2 but across different values of α_1 and N; 'Overall average' in columns 3 - 6 is the overall average for each N for different values of α_1 and α_2 while 'Overall average' in the last two column is the overall average for all the cases in the entire table.

Table B3: Rejection Rate, error ∼ t(2)

α_2	α_1	N=100	N=200	N=400	N=800	Raverage	Saverage
	0.1	0.055	0.052	0.049	0.048	0.051	
	0.3	0.055	0.053	0.051	0.051	0.053	
0.1	0.5	0.063	0.065	0.061	0.058	0.062	0.060
	0.7	0.068	0.064	0.065	0.060	0.064	
	0.9	0.072	0.076	0.074	0.070	0.073	
Caverage		0.062	0.062	0.06	0.057	0.060	0.060
	0.1	0.060	0.054	0.056	0.051	0.055	
	0.3	0.072	0.068	0.061	0.064	0.067	
0.3	0.5	0.087	0.086	0.084	0.082	0.085	0.088
	0.7	0.112	0.106	0.102	0.101	0.105	
	0.9	0.139	0.129	0.128	0.126	0.130	
Caverage		0.094	0.089	0.086	0.085	0.088	0.099
	0.1	0.062	0.065	0.058	0.061	0.061	
	0.3	0.083	0.084	0.080	0.080	0.082	
0.5	0.5	0.119	0.115	0.107	0.110	0.113	0.126
	0.7	0.161	0.156	0.151	0.151	0.155	
	0.9	0.222	0.216	0.217	0.212	0.217	
Caverage		0.129	0.127	0.123	0.123	0.126	0.126
	0.1	0.071	0.068	0.064	0.063	0.066	
	0.3	0.106	0.100	0.101	0.100	0.102	
0.7	0.5	0.161	0.159	0.153	0.155	0.157	0.177
	0.7	0.234	0.234	0.230	0.226	0.231	
	0.9	0.327	0.329	0.328	0.327	0.328	
Caverage		0.180	0.179	0.174	0.174	0.177	0.177
	0.1	0.073	0.075	0.074	0.073	0.074	
	0.3	0.134	0.129	0.122	0.128	0.128	
0.9	0.5	0.210	0.220	0.220	0.210	0.215	0.251
	0.7	0.335	0.332	0.339	0.328	0.334	
	0.9	0.498	0.513	0.507	0.508	0.506	
Caverage		0.250	0.254	0.252	0.250	0.251	0.251
Overall average		0.143	0.142	0.139	0.138	0.140	0.140

Note: 'Raverage' (stands for row average) is the average rejection rate for the same value of α_1 and α_2 but across different values of N ; 'Caverage' (stands for column average) is the average rejection rate for the same value of N and α_2 but across different values of α_1; 'Saverage' (stands for Situation average) is the average rejection rate for the same value of α_2 but across different values of α_1 and N; 'Overall average' in columns 3 - 6 is the overall average for each N for different values of α_1 and α_2 while 'Overall average' in the last two column is the overall average for all the cases in the entire table.

Table B4: Rejection Rate, error ∼ t(1)

α_2	α_1	N=100	N=200	N=400	N=800	Raverage	Saverage
0.1	0.1	0.046	0.039	0.028	0.025	0.035	0.050
	0.3	0.048	0.042	0.035	0.031	0.039	
	0.5	0.062	0.054	0.042	0.035	0.048	
	0.7	0.067	0.057	0.049	0.042	0.054	
	0.9	0.080	0.077	0.070	0.061	0.072	
Caverage		0.061	0.054	0.045	0.039	0.050	0.050
0.3	0.1	0.054	0.044	0.035	0.030	0.041	0.063
	0.3	0.057	0.051	0.042	0.032	0.045	
	0.5	0.072	0.062	0.048	0.041	0.056	
	0.7	0.090	0.076	0.065	0.047	0.070	
	0.9	0.120	0.114	0.097	0.079	0.102	
Caverage		0.079	0.069	0.057	0.048	0.063	0.063
0.5	0.1	0.059	0.051	0.041	0.033	0.046	0.084
	0.3	0.070	0.059	0.047	0.040	0.054	
	0.5	0.086	0.073	0.064	0.053	0.069	
	0.7	0.125	0.107	0.083	0.069	0.096	
	0.9	0.193	0.164	0.143	0.114	0.153	
Caverage		0.107	0.091	0.076	0.062	0.084	0.084
0.7	0.1	0.068	0.065	0.053	0.040	0.056	0.119
	0.3	0.092	0.078	0.067	0.053	0.072	
	0.5	0.127	0.106	0.085	0.071	0.097	
	0.7	0.173	0.149	0.126	0.098	0.137	
	0.9	0.290	0.256	0.213	0.176	0.234	
Caverage		0.150	0.131	0.109	0.088	0.119	0.119
0.9	0.1	0.077	0.077	0.064	0.060	0.070	0.194
	0.3	0.121	0.111	0.095	0.082	0.102	
	0.5	0.193	0.168	0.138	0.122	0.155	
	0.7	0.289	0.258	0.212	0.178	0.234	
	0.9	0.479	0.435	0.385	0.329	0.407	
Caverage		0.232	0.210	0.179	0.154	0.194	0.194
Overall average		0.126	0.111	0.093	0.078	0.102	0.102

Note: 'Raverage' (stands for row average) is the average rejection rate for the same value of α_1 and α_2 but across different values of N ; 'Caverage' (stands for column average) is the average rejection rate for the same value of N and α_2 but across different values of α_1; 'Saverage' (stands for Situation average) is the average rejection rate for the same value of α_2 but across different values of α_1 and N; 'Overall average' in columns 3 - 6 is the overall average for each N for different values of α_1 and α_2 while 'Overall average' in the last two column is the overall average for all the cases in the entire table.

C The rejection rate for $(\alpha_1, \alpha_2) \in (A^-, A^-)$

Table C1: Rejection Rate, error $\sim$ N(0,1)

α_2	α_1	N=100	N=200	N=400	N=800	Raverage	Saverage
	-0.1	0.055	0.051	0.050	0.053	0.052	
	-0.3	0.055	0.058	0.058	0.056	0.057	
-0.1	-0.5	0.067	0.064	0.061	0.063	0.064	0.063
	-0.7	0.067	0.068	0.066	0.069	0.068	
	-0.9	0.070	0.075	0.072	0.074	0.073	
Caverage		0.063	0.063	0.061	0.063	0.063	0.063
	-0.1	0.058	0.061	0.066	0.062	0.062	
	-0.3	0.074	0.074	0.073	0.075	0.074	
-0.3	-0.5	0.093	0.093	0.097	0.093	0.094	0.096
	-0.7	0.105	0.116	0.111	0.109	0.110	
	-0.9	0.135	0.140	0.137	0.144	0.139	
Caverage		0.093	0.097	0.097	0.097	0.096	0.096
	-0.1	0.066	0.064	0.063	0.063	0.064	
	-0.3	0.092	0.092	0.103	0.089	0.094	
-0.5	-0.5	0.127	0.128	0.128	0.131	0.128	0.137
	-0.7	0.170	0.176	0.179	0.177	0.175	
	-0.9	0.228	0.225	0.224	0.225	0.225	
Caverage		0.137	0.137	0.139	0.137	0.137	0.137
	-0.1	0.069	0.074	0.067	0.070	0.070	
	-0.3	0.115	0.111	0.107	0.109	0.110	
-0.7	-0.5	0.166	0.168	0.166	0.178	0.170	0.190
	-0.7	0.249	0.251	0.248	0.248	0.249	
	-0.9	0.348	0.349	0.358	0.356	0.353	
Caverage		0.190	0.191	0.189	0.192	0.190	0.190
	-0.1	0.073	0.071	0.073	0.071	0.072	
	-0.3	0.140	0.132	0.137	0.135	0.136	
-0.9	-0.5	0.225	0.226	0.229	0.231	0.228	0.262
	-0.7	0.344	0.346	0.350	0.347	0.347	
	-0.9	0.526	0.523	0.527	0.531	0.527	
Caverage		0.261	0.260	0.263	0.263	0.262	0.262
Overall average		0.149	0.149	0.150	0.150	0.150	0.150

Note: 'Raverage' (stands for row average) is the average rejection rate for the same value of α_1 and α_2 but across different values of N ; 'Caverage' (stands for column average) is the average rejection rate for the same value of N and α_2 but across different values of α_1; 'Saverage' (stands for Situation average) is the average rejection rate for the same value of α_2 but across different values of α_1 and N; 'Overall average' in columns 3 - 6 is the overall average for each N for different values of α_1 and α_2 while 'Overall average' in the last two column is the overall average for all the cases in the entire table.

Table C2: Rejection Rate, error ∼ t(5)

α_2	α_1	N=100	N=200	N=400	N=800	Raverage	Saverage
	-0.1	0.053	0.053	0.051	0.053	0.053	
	-0.3	0.058	0.060	0.055	0.051	0.056	
-0.1	-0.5	0.062	0.062	0.062	0.063	0.062	0.063
	-0.7	0.065	0.068	0.070	0.067	0.067	
	-0.9	0.075	0.076	0.070	0.076	0.074	
Caverage		0.063	0.064	0.062	0.062	0.063	0.063
	-0.1	0.057	0.058	0.059	0.060	0.058	
	-0.3	0.076	0.073	0.074	0.073	0.074	
-0.3	-0.5	0.092	0.095	0.089	0.089	0.091	0.094
	-0.7	0.108	0.111	0.110	0.107	0.109	
	-0.9	0.134	0.133	0.139	0.141	0.137	
Caverage		0.093	0.094	0.094	0.094	0.094	0.094
	-0.1	0.064	0.062	0.064	0.061	0.063	
	-0.3	0.091	0.094	0.091	0.089	0.091	
-0.5	-0.5	0.127	0.130	0.133	0.128	0.129	0.136
	-0.7	0.173	0.174	0.173	0.171	0.172	
	-0.9	0.222	0.219	0.226	0.228	0.223	
Caverage		0.135	0.136	0.137	0.135	0.136	0.136
	-0.1	0.065	0.065	0.072	0.067	0.067	
	-0.3	0.116	0.108	0.116	0.111	0.113	
-0.7	-0.5	0.165	0.175	0.171	0.173	0.171	0.190
	-0.7	0.251	0.245	0.256	0.255	0.252	
	-0.9	0.339	0.347	0.357	0.351	0.349	
Caverage		0.187	0.188	0.194	0.191	0.190	0.190
	-0.1	0.077	0.073	0.076	0.077	0.076	
	-0.3	0.138	0.142	0.139	0.132	0.138	
-0.9	-0.5	0.225	0.219	0.233	0.228	0.226	0.263
	-0.7	0.352	0.349	0.353	0.354	0.352	
	-0.9	0.514	0.522	0.521	0.528	0.521	
Caverage		0.261	0.261	0.265	0.264	0.263	0.263
Overall average		0.148	0.148	0.150	0.149	0.149	0.149

Note: 'Raverage' (stands for row average) is the average rejection rate for the same value of α_1 and α_2 but across different values of N ; 'Caverage' (stands for column average) is the average rejection rate for the same value of N and α_2 but across different values of α_1; 'Saverage' (stands for Situation average) is the average rejection rate for the same value of α_2 but across different values of α_1 and N; 'Overall average' in columns 3 - 6 is the overall average for each N for different values of α_1 and α_2 while 'Overall average' in the last two column is the overall average for all the cases in the entire table.

Table C3: Rejection Rate, error $\sim$ t(2)

α_2	α_1	N=100	N=200	N=400	N=800	Raverage	Saverage
	-0.1	0.054	0.053	0.048	0.046	0.050	
	-0.3	0.061	0.055	0.054	0.054	0.056	
-0.1	-0.5	0.062	0.063	0.063	0.058	0.061	0.061
	-0.7	0.067	0.066	0.061	0.063	0.064	
	-0.9	0.074	0.075	0.074	0.071	0.073	
Caverage		0.064	0.062	0.060	0.058	0.061	0.061
	-0.1	0.060	0.053	0.057	0.051	0.055	
	-0.3	0.070	0.072	0.063	0.061	0.066	
-0.3	-0.5	0.086	0.083	0.080	0.078	0.082	0.088
	-0.7	0.111	0.109	0.104	0.096	0.105	
	-0.9	0.137	0.134	0.128	0.129	0.132	
Caverage		0.093	0.090	0.086	0.083	0.088	0.088
	-0.1	0.064	0.061	0.057	0.057	0.060	
	-0.3	0.087	0.083	0.083	0.083	0.084	
-0.5	-0.5	0.119	0.114	0.107	0.106	0.111	0.125
	-0.7	0.163	0.155	0.154	0.152	0.156	
	-0.9	0.216	0.223	0.213	0.211	0.216	
Caverage		0.130	0.127	0.123	0.122	0.125	0.125
	-0.1	0.071	0.065	0.065	0.061	0.066	
	-0.3	0.102	0.103	0.098	0.098	0.100	
-0.7	-0.5	0.159	0.153	0.154	0.147	0.153	0.176
	-0.7	0.235	0.230	0.224	0.218	0.227	
	-0.9	0.341	0.334	0.332	0.332	0.335	
Caverage		0.182	0.177	0.175	0.171	0.176	0.176
	-0.1	0.071	0.074	0.070	0.073	0.072	
	-0.3	0.137	0.135	0.128	0.132	0.133	
-0.9	-0.5	0.218	0.215	0.219	0.202	0.213	0.253
	-0.7	0.339	0.335	0.335	0.327	0.334	
	-0.9	0.517	0.512	0.512	0.503	0.511	
Caverage		0.256	0.254	0.253	0.247	0.253	0.253
Overall average		0.145	0.142	0.139	0.136	0.141	0.141

Note: 'Raverage' (stands for row average) is the average rejection rate for the same value of α_1 and α_2 but across different values of N ; 'Caverage' (stands for column average) is the average rejection rate for the same value of N and α_2 but across different values of α_1; 'Saverage' (stands for Situation average) is the average rejection rate for the same value of α_2 but across different values of α_1 and N; 'Overall average' in columns 3 - 6 is the overall average for each N for different values of α_1 and α_2 while 'Overall average' in the last two column is the overall average for all the cases in the entire table.

Table C4: Rejection Rate, error ∼ t(1)

α_2	α_1	N=100	N=200	N=400	N=800	Raverage	Saverage
	-0.1	0.047	0.040	0.029	0.026	0.035	
	-0.3	0.051	0.043	0.035	0.031	0.040	
-0.1	-0.5	0.058	0.051	0.043	0.036	0.047	0.050
	-0.7	0.069	0.058	0.052	0.042	0.055	
	-0.9	0.079	0.076	0.072	0.063	0.073	
Caverage		0.061	0.054	0.046	0.039	0.050	0.050
	-0.1	0.054	0.045	0.035	0.029	0.041	
	-0.3	0.058	0.053	0.042	0.034	0.047	
-0.3	-0.5	0.074	0.061	0.049	0.044	0.057	0.065
	-0.7	0.093	0.081	0.067	0.049	0.072	
	-0.9	0.129	0.115	0.098	0.083	0.106	
Caverage		0.082	0.071	0.058	0.048	0.065	0.065
	-0.1	0.061	0.048	0.041	0.033	0.046	
	-0.3	0.074	0.060	0.049	0.040	0.056	
-0.5	-0.5	0.095	0.076	0.069	0.053	0.073	0.085
	-0.7	0.127	0.108	0.085	0.069	0.097	
	-0.9	0.193	0.167	0.141	0.113	0.153	
Caverage		0.110	0.092	0.077	0.061	0.085	0.085
	-0.1	0.067	0.064	0.052	0.038	0.055	
	-0.3	0.090	0.080	0.066	0.057	0.073	
-0.7	-0.5	0.132	0.109	0.086	0.072	0.100	0.120
	-0.7	0.180	0.150	0.127	0.095	0.138	
	-0.9	0.299	0.252	0.214	0.178	0.236	
Caverage		0.153	0.131	0.109	0.088	0.120	0.120
	-0.1	0.080	0.077	0.065	0.060	0.070	
	-0.3	0.127	0.116	0.100	0.082	0.106	
-0.9	-0.5	0.189	0.173	0.139	0.119	0.155	0.192
	-0.7	0.288	0.249	0.214	0.181	0.233	
	-0.9	0.465	0.426	0.371	0.314	0.394	
Caverage		0.230	0.208	0.178	0.151	0.192	0.192
Overall average		0.127	0.111	0.094	0.078	0.102	0.102

Note: 'Raverage' (stands for row average) is the average rejection rate for the same value of α_1 and α_2 but across different values of N ; 'Caverage' (stands for column average) is the average rejection rate for the same value of N and α_2 but across different values of α_1; 'Saverage' (stands for Situation average) is the average rejection rate for the same value of α_2 but across different values of α_1 and N; 'Overall average' in columns 3 - 6 is the overall average for each N for different values of α_1 and α_2 while 'Overall average' in the last two column is the overall average for all the cases in the entire table.

D The rejection rate for $(\alpha_1, \alpha_2) \in (A^+, A^-)$

Table D1: Rejection Rate, error ~ N(0,1)

α_2	α_1	N=100	N=200	N=400	N=800	Raverage	Saverage
	0.1	0.052	0.046	0.045	0.049	0.048	
	0.3	0.045	0.041	0.045	0.044	0.043	
-0.1	0.5	0.039	0.042	0.038	0.042	0.040	0.040
	0.7	0.037	0.037	0.035	0.035	0.036	
	0.9	0.035	0.032	0.033	0.031	0.033	
Caverage		0.042	0.040	0.039	0.040	0.040	0.040
	0.1	0.045	0.046	0.043	0.049	0.046	
	0.3	0.035	0.032	0.033	0.032	0.033	
-0.3	0.5	0.024	0.024	0.023	0.023	0.024	0.026
	0.7	0.016	0.016	0.015	0.015	0.015	
	0.9	0.010	0.010	0.010	0.011	0.010	
Caverage		0.026	0.026	0.025	0.026	0.026	0.026
	0.1	0.043	0.041	0.037	0.040	0.040	
	0.3	0.026	0.023	0.027	0.024	0.025	
-0.5	0.5	0.013	0.012	0.014	0.013	0.013	0.017
	0.7	0.007	0.007	0.005	0.004	0.006	
	0.9	0.003	0.002	0.002	0.002	0.002	
Caverage		0.018	0.017	0.017	0.017	0.017	0.017
	0.1	0.039	0.038	0.037	0.035	0.037	
	0.3	0.018	0.017	0.017	0.016	0.017	
-0.7	0.5	0.007	0.005	0.005	0.005	0.006	0.012
	0.7	0.001	0.001	0.001	0.000	0.001	
	0.9	0.000	0.000	0.000	0.000	0.000	
Caverage		0.013	0.012	0.012	0.011	0.012	0.012
	0.1	0.034	0.033	0.033	0.033	0.033	
	0.3	0.010	0.009	0.010	0.009	0.009	
-0.9	0.5	0.003	0.002	0.002	0.001	0.002	0.009
	0.7	0.000	0.000	0.000	0.000	0.000	
	0.9	0.000	0.000	0.000	0.000	0.000	
Caverage		0.009	0.009	0.009	0.009	0.009	0.009
Overall average		0.022	0.021	0.020	0.020	0.021	0.021

Note: 'Raverage' (stands for row average) is the average rejection rate for the same value of α_1 and α_2 but across different values of N ; 'Caverage' (stands for column average) is the average rejection rate for the same value of N and α_2 but across different values of α_1; 'Saverage' (stands for Situation average) is the average rejection rate for the same value of α_2 but across different values of α_1 and N; 'Overall average' in columns 3 - 6 is the overall average for each N for different values of α_1 and α_2 while 'Overall average' in the last two column is the overall average for all the cases in the entire table.

Table D2: Rejection Rate, error $\sim$ t(5)

α_2	α_1	N=100	N=200	N=400	N=800	Raverage	Saverage
	0.1	0.049	0.050	0.047	0.047	0.048	
	0.3	0.045	0.046	0.042	0.043	0.044	
-0.1	0.5	0.043	0.039	0.044	0.038	0.041	0.040
	0.7	0.037	0.037	0.035	0.035	0.036	
	0.9	0.031	0.031	0.035	0.034	0.033	
Caverage		0.041	0.041	0.041	0.039	0.040	0.040
	0.1	0.044	0.046	0.044	0.045	0.045	
	0.3	0.034	0.033	0.034	0.029	0.032	
-0.3	0.5	0.024	0.024	0.023	0.024	0.024	0.025
	0.7	0.019	0.018	0.014	0.014	0.016	
	0.9	0.010	0.011	0.011	0.010	0.010	
Caverage		0.026	0.026	0.025	0.024	0.025	0.025
	0.1	0.043	0.041	0.040	0.036	0.040	
	0.3	0.025	0.023	0.022	0.022	0.023	
-0.5	0.5	0.014	0.013	0.012	0.012	0.013	0.017
	0.7	0.006	0.005	0.005	0.006	0.005	
	0.9	0.002	0.002	0.002	0.002	0.002	
Caverage		0.018	0.017	0.016	0.016	0.017	0.017
	0.1	0.035	0.036	0.038	0.035	0.036	
	0.3	0.017	0.015	0.015	0.017	0.016	
-0.7	0.5	0.006	0.006	0.004	0.006	0.005	0.012
	0.7	0.001	0.001	0.001	0.000	0.001	
	0.9	0.000	0.000	0.000	0.000	0.000	
Caverage		0.038	0.036	0.035	0.034	0.012	0.012
	0.1	0.011	0.011	0.010	0.009	0.010	
	0.3	0.003	0.003	0.002	0.001	0.002	
-0.9	0.5	0.001	0.000	0.000	0.000	0.000	0.002
	0.7	0.000	0.000	0.000	0.000	0.000	
	0.9	0.000	0.000	0.000	0.000	0.000	
Caverage		0.003	0.003	0.002	0.002	0.002	0.002
Overall average		0.021	0.021	0.021	0.020	0.021	0.021

Note: 'Raverage' (stands for row average) is the average rejection rate for the same value of α_1 and α_2 but across different values of N ; 'Caverage' (stands for column average) is the average rejection rate for the same value of N and α_2 but across different values of α_1; 'Saverage' (stands for Situation average) is the average rejection rate for the same value of α_2 but across different values of α_1 and N; 'Overall average' in columns 3 - 6 is the overall average for each N for different values of α_1 and α_2 while 'Overall average' in the last two column is the overall average for all the cases in the entire table.

Table D3: Rejection Rate, error ~ t(2)

α_2	α_1	N=100	N=200	N=400	N=800	Raverage	Saverage
	0.1	0.051	0.049	0.049	0.043	0.048	
	0.3	0.044	0.044	0.045	0.043	0.044	
-0.1	0.5	0.044	0.044	0.043	0.042	0.043	0.042
	0.7	0.040	0.037	0.037	0.037	0.038	
	0.9	0.036	0.038	0.031	0.035	0.035	
Caverage		0.043	0.042	0.041	0.040	0.042	0.042
	0.1	0.049	0.045	0.044	0.042	0.045	
	0.3	0.039	0.037	0.034	0.035	0.036	
-0.3	0.5	0.029	0.030	0.024	0.025	0.027	0.028
	0.7	0.021	0.025	0.019	0.020	0.021	
	0.9	0.013	0.013	0.010	0.015	0.012	
Caverage		0.030	0.030	0.026	0.027	0.028	0.028
	0.1	0.046	0.044	0.043	0.040	0.043	
	0.3	0.029	0.029	0.025	0.030	0.028	
-0.5	0.5	0.020	0.017	0.014	0.019	0.017	0.020
	0.7	0.011	0.011	0.006	0.012	0.010	
	0.9	0.004	0.004	0.002	0.006	0.004	
Caverage		0.022	0.021	0.018	0.021	0.020	0.020
	0.1	0.046	0.041	0.035	0.038	0.040	
	0.3	0.023	0.023	0.017	0.022	0.021	
-0.7	0.5	0.010	0.011	0.006	0.012	0.010	0.015
	0.7	0.004	0.005	0.001	0.006	0.004	
	0.9	0.001	0.000	0.000	0.001	0.001	
Caverage		0.017	0.016	0.012	0.016	0.015	0.015
	0.1	0.033	0.034	0.038	0.035	0.035	
	0.3	0.014	0.013	0.011	0.017	0.014	
-0.9	0.5	0.005	0.005	0.003	0.004	0.004	0.011
	0.7	0.001	0.000	0.001	0.002	0.001	
	0.9	0.000	0.000	0.000	0.000	0.000	
Caverage		0.011	0.011	0.010	0.012	0.011	0.011
Overall average		0.024	0.024	0.021	0.023	0.023	0.023

Note: 'Raverage' (stands for row average) is the average rejection rate for the same value of α_1 and α_2 but across different values of N ; 'Caverage' (stands for column average) is the average rejection rate for the same value of N and α_2 but across different values of α_1; 'Saverage' (stands for Situation average) is the average rejection rate for the same value of α_2 but across different values of α_1 and N; 'Overall average' in columns 3 - 6 is the overall average for each N for different values of α_1 and α_2 while 'Overall average' in the last two column is the overall average for all the cases in the entire table.

Table D4: Rejection Rate, error $\sim$ t(1)

α_2	α_1	N=100	N=200	N=400	N=800	Raverage	Saverage
	0.1	0.045	0.039	0.028	0.025	0.034	
	0.3	0.044	0.039	0.032	0.028	0.036	
-0.1	0.5	0.051	0.044	0.035	0.030	0.040	0.040
	0.7	0.049	0.045	0.034	0.035	0.042	
	0.9	0.047	0.048	0.049	0.045	0.047	
Caverage		0.047	0.043	0.037	0.033	0.040	0.040
	0.1	0.049	0.041	0.033	0.027	0.037	
	0.3	0.041	0.040	0.032	0.024	0.034	
-0.3	0.5	0.040	0.036	0.030	0.026	0.033	0.033
	0.7	0.039	0.035	0.032	0.025	0.033	
	0.9	0.024	0.031	0.032	0.028	0.029	
Caverage		0.039	0.037	0.032	0.026	0.033	0.033
	0.1	0.050	0.041	0.037	0.029	0.039	
	0.3	0.039	0.037	0.031	0.025	0.033	
-0.5	0.5	0.034	0.030	0.027	0.024	0.029	0.028
	0.7	0.026	0.026	0.022	0.022	0.024	
	0.9	0.009	0.014	0.019	0.019	0.015	
Caverage		0.031	0.030	0.027	0.024	0.028	0.028
	0.1	0.050	0.051	0.0441	0.031	0.043	
	0.3	0.036	0.036	0.029	0.027	0.032	
-0.7	0.5	0.028	0.026	0.024	0.021	0.025	0.024
	0.7	0.014	0.016	0.016	0.015	0.015	
	0.9	0.003	0.004	0.007	0.011	0.006	
Caverage		0.026	0.027	0.024	0.021	0.024	0.024
	0.1	0.051	0.052	0.046	0.045	0.048	
	0.3	0.027	0.030	0.030	0.030	0.029	
-0.9	0.5	0.011	0.015	0.019	0.020	0.016	0.020
	0.7	0.003	0.005	0.009	0.010	0.007	
	0.9	0.001	0.000	0.001	0.002	0.001	
Caverage		0.019	0.020	0.021	0.021	0.028	0.020
Overall average		0.033	0.031	0.028	0.025	0.029	0.029

Note: 'Raverage' (stands for row average) is the average rejection rate for the same value of α_1 and α_2 but across different values of N ; 'Caverage' (stands for column average) is the average rejection rate for the same value of N and α_2 but across different values of α_1; 'Saverage' (stands for Situation average) is the average rejection rate for the same value of α_2 but across different values of α_1 and N; 'Overall average' in columns 3 - 6 is the overall average for each N for different values of α_1 and α_2 while 'Overall average' in the last two column is the overall average for all the cases in the entire table.

E The rejection rate for $(\alpha_1, \alpha_2) \in (A^-, A^+)$

Table E1: Rejection Rate, error $\sim$ N(0,1)

α_2	α_1	N=100	N=200	N=400	N=800	Raverage	Saverage
	-0.1	0.051	0.046	0.046	0.049	0.048	
	-0.3	0.045	0.044	0.043	0.044	0.044	
0.1	-0.5	0.043	0.042	0.040	0.042	0.042	0.040
	-0.7	0.037	0.034	0.035	0.035	0.035	
	-0.9	0.032	0.033	0.031	0.031	0.032	
Caverage		0.042	0.040	0.039	0.040	0.040	0.040
	-0.1	0.045	0.044	0.042	0.049	0.045	
	-0.3	0.033	0.035	0.030	0.032	0.033	
0.3	-0.5	0.025	0.023	0.024	0.023	0.024	0.026
	-0.7	0.016	0.018	0.016	0.015	0.016	
	-0.9	0.011	0.012	0.011	0.011	0.011	
Caverage		0.026	0.026	0.025	0.026	0.026	0.026
	-0.1	0.042	0.038	0.041	0.040	0.040	
	-0.3	0.026	0.022	0.024	0.024	0.024	
0.5	-0.5	0.012	0.011	0.011	0.013	0.012	0.017
	-0.7	0.005	0.006	0.005	0.004	0.005	
	-0.9	0.003	0.003	0.001	0.002	0.002	
Caverage		0.018	0.016	0.016	0.017	0.017	0.017
	-0.1	0.035	0.036	0.036	0.035	0.035	
	-0.3	0.018	0.016	0.018	0.016	0.017	
0.7	-0.5	0.008	0.007	0.004	0.005	0.006	0.012
	-0.7	0.002	0.002	0.001	0.000	0.001	
	-0.9	0.000	0.000	0.000	0.000	0.000	
Caverage		0.013	0.012	0.012	0.011	0.012	0.012
	-0.1	0.034	0.033	0.030	0.033	0.032	
	-0.3	0.011	0.012	0.008	0.009	0.010	
0.9	-0.5	0.003	0.001	0.002	0.001	0.002	0.009
	-0.7	0.000	0.000	0.000	0.000	0.000	
	-0.9	0.000	0.000	0.000	0.000	0.000	
Caverage		0.009	0.009	0.008	0.009	0.009	0.009
Overall average		0.021	0.021	0.020	0.020	0.021	0.021

Note: 'Raverage' (stands for row average) is the average rejection rate for the same value of α_1 and α_2 but across different values of N ; 'Caverage' (stands for column average) is the average rejection rate for the same value of N and α_2 but across different values of α_1; 'Saverage' (stands for Situation average) is the average rejection rate for the same value of α_2 but across different values of α_1 and N; 'Overall average' in columns 3 - 6 is the overall average for each N for different values of α_1 and α_2 while 'Overall average' in the last two column is the overall average for all the cases in the entire table.

Table E2: Rejection Rate, error ∼ t(5)

α_2	α_1	N=100	N=200	N=400	N=800	Raverage	Saverage
	-0.1	0.050	0.050	0.049	0.049	0.049	
	-0.3	0.048	0.046	0.042	0.039	0.044	
0.1	-0.5	0.042	0.040	0.041	0.039	0.041	0.041
	-0.7	0.035	0.037	0.039	0.038	0.037	
	-0.9	0.035	0.034	0.032	0.032	0.033	
Caverage		0.042	0.041	0.041	0.039	0.041	0.041
	-0.1	0.047	0.046	0.045	0.047	0.046	
	-0.3	0.035	0.032	0.035	0.032	0.033	
0.3	-0.5	0.024	0.026	0.020	0.023	0.023	0.026
	-0.7	0.016	0.015	0.017	0.013	0.015	
	-0.9	0.012	0.011	0.010	0.010	0.011	
Caverage		0.027	0.026	0.026	0.025	0.026	0.026
	-0.1	0.043	0.038	0.042	0.041	0.041	
	-0.3	0.024	0.027	0.024	0.021	0.024	
0.5	-0.5	0.014	0.014	0.013	0.012	0.013	0.017
	-0.7	0.007	0.006	0.004	0.004	0.005	
	-0.9	0.003	0.002	0.002	0.002	0.002	
Caverage		0.018	0.017	0.017	0.016	0.017	0.017
	-0.1	0.038	0.036	0.036	0.034	0.036	
	-0.3	0.018	0.014	0.015	0.015	0.016	
0.7	-0.5	0.006	0.006	0.006	0.006	0.006	0.012
	-0.7	0.002	0.002	0.001	0.001	0.001	
	-0.9	0.000	0.000	0.000	0.000	0.000	
Caverage		0.013	0.011	0.012	0.011	0.012	0.012
	-0.1	0.036	0.033	0.036	0.033	0.035	
	-0.3	0.012	0.011	0.012	0.011	0.011	
0.9	-0.5	0.003	0.002	0.001	0.002	0.002	0.010
	-0.7	0.001	0.000	0.000	0.000	0.000	
	-0.9	0.000	0.000	0.000	0.000	0.000	
Caverage		0.010	0.009	0.010	0.009	0.010	0.010
Overall average		0.022	0.021	0.021	0.020	0.021	0.021

Note: 'Raverage' (stands for row average) is the average rejection rate for the same value of α_1 and α_2 but across different values of N ; 'Caverage' (stands for column average) is the average rejection rate for the same value of N and α_2 but across different values of α_1; 'Saverage' (stands for Situation average) is the average rejection rate for the same value of α_2 but across different values of α_1 and N; 'Overall average' in columns 3 - 6 is the overall average for each N for different values of α_1 and α_2 while 'Overall average' in the last two column is the overall average for all the cases in the entire table.

Table E3: Rejection Rate, error $\sim$ t(2)

α_2	α_1	N=100	N=200	N=400	N=800	Raverage	Saverage
	-0.1	0.051	0.043	0.043	0.043	0.045	
	-0.3	0.044	0.043	0.044	0.045	0.044	
0.1	-0.5	0.041	0.042	0.046	0.042	0.043	0.042
	-0.7	0.036	0.037	0.038	0.036	0.037	
	-0.9	0.050	0.035	0.039	0.035	0.040	
Caverage		0.044	0.040	0.042	0.040	0.042	0.042
	-0.1	0.038	0.042	0.046	0.042	0.042	
	-0.3	0.031	0.035	0.035	0.033	0.034	
0.3	-0.5	0.027	0.025	0.026	0.029	0.027	0.029
	-0.7	0.015	0.020	0.021	0.020	0.019	
	-0.9	0.042	0.015	0.015	0.014	0.021	
Caverage		0.031	0.027	0.028	0.028	0.029	0.029
	-0.1	0.028	0.040	0.041	0.044	0.038	
	-0.3	0.022	0.030	0.026	0.030	0.027	
0.5	-0.5	0.011	0.019	0.018	0.018	0.016	0.021
	-0.7	0.003	0.012	0.012	0.011	0.009	
	-0.9	0.042	0.006	0.004	0.005	0.014	
Caverage		0.021	0.021	0.020	0.021	0.021	0.021
	-0.1	0.023	0.038	0.036	0.039	0.034	
	-0.3	0.011	0.022	0.021	0.021	0.019	
0.7	-0.5	0.003	0.012	0.011	0.011	0.009	0.015
	-0.7	0.001	0.006	0.005	0.004	0.004	
	-0.9	0.034	0.001	0.001	0.001	0.009	
Caverage		0.014	0.016	0.015	0.015	0.015	0.015
	-0.1	0.033	0.034	0.038	0.035	0.035	
	-0.3	0.013	0.035	0.037	0.034	0.030	
0.9	-0.5	0.004	0.017	0.013	0.015	0.012	0.016
	-0.7	0.001	0.004	0.005	0.005	0.004	
	-0.9	0.000	0.002	0.001	0.001	0.001	
Caverage		0.000	0.000	0.000	0.000	0.016	0.016
Caverage		0.023	0.023	0.023	0.023	0.023	0.023

Note: 'Raverage' (stands for row average) is the average rejection rate for the same value of α_1 and α_2 but across different values of N ; 'Caverage' (stands for column average) is the average rejection rate for the same value of N and α_2 but across different values of α_1; 'Saverage' (stands for Situation average) is the average rejection rate for the same value of α_2 but across different values of α_1 and N; 'Overall average' in columns 3 - 6 is the overall average for each N for different values of α_1 and α_2 while 'Overall average' in the last two column is the overall average for all the cases in the entire table.

Table E4: Rejection Rate, error ∼ t(1)

α_2	α_1	N=100	N=200	N=400	N=800	Raverage	Saverage
	-0.1	0.045	0.039	0.028	0.025	0.034	
	-0.3	0.045	0.039	0.032	0.028	0.036	
0.1	-0.5	0.049	0.044	0.037	0.032	0.040	0.041
	-0.7	0.052	0.046	0.042	0.035	0.044	
	-0.9	0.049	0.051	0.049	0.047	0.049	
Caverage		0.048	0.044	0.038	0.033	0.041	0.041
	-0.1	0.048	0.041	0.033	0.027	0.037	
	-0.3	0.042	0.038	0.032	0.023	0.034	
0.3	-0.5	0.041	0.035	0.031	0.026	0.033	0.033
	-0.7	0.039	0.037	0.033	0.024	0.033	
	-0.9	0.027	0.032	0.034	0.028	0.030	
Caverage		0.039	0.036	0.033	0.026	0.033	0.033
	-0.1	0.050	0.044	0.036	0.027	0.039	
	-0.3	0.040	0.036	0.030	0.025	0.032	
0.5	-0.5	0.034	0.029	0.027	0.024	0.029	0.028
	-0.7	0.026	0.027	0.023	0.021	0.024	
	-0.9	0.010	0.015	0.022	0.020	0.017	
Caverage		0.032	0.030	0.027	0.024	0.028	0.028
	-0.1	0.050	0.050	0.043	0.031	0.043	
	-0.3	0.035	0.036	0.029	0.027	0.032	
0.7	-0.5	0.027	0.025	0.024	0.020	0.024	0.024
	-0.7	0.014	0.016	0.017	0.013	0.015	
	-0.9	0.003	0.004	0.007	0.012	0.007	
Caverage		0.026	0.026	0.024	0.021	0.024	0.024
	-0.1	0.045	0.051	0.045	0.044	0.046	
	-0.3	0.025	0.030	0.029	0.030	0.029	
0.9	-0.5	0.010	0.014	0.019	0.020	0.016	0.020
	-0.7	0.003	0.003	0.010	0.009	0.006	
	-0.9	0.001	0.001	0.001	0.003	0.001	
Caverage		0.017	0.020	0.021	0.021	0.020	0.020
Overall average		0.032	0.031	0.028	0.025	0.029	0.029

Note: 'Raverage' (stands for row average) is the average rejection rate for the same value of α_1 and α_2 but across different values of N ; 'Caverage' (stands for column average) is the average rejection rate for the same value of N and α_2 but across different values of α_1; 'Saverage' (stands for Situation average) is the average rejection rate for the same value of α_2 but across different values of α_1 and N; 'Overall average' in columns 3 - 6 is the overall average for each N for different values of α_1 and α_2 while 'Overall average' in the last two column is the overall average for all the cases in the entire table.

F Comparison of rejection rate before and after transformation for $(\alpha_1, \alpha_2) \in (A^+, A^+)$

Table F1: Comparison of rejection rates before and after transformation in all cases, error ∼ N(0,1)

α_2	α_1	N=100	N=200	N=400	N=800	Raverage	Saverage
	0.1	0.056(0.058)	0.052(0.053)	0.052(0.049)	0.055(0.053)	1.31%	
	0.3	0.059(0.060)	0.056(0.055)	0.060(0.055)	0.054(0.050)	3.87%	
0.1	0.5	0.063(0.062)	0.063(0.059)	0.061(0.054)	0.066(0.056)	8.85%	9.17%
	0.7	0.069(0.065)	0.068(0.060)	0.067(0.055)	0.068(0.054)	14.06%	
	0.9	0.073(0.067)	0.075(0.062)	0.074(0.059)	0.071(0.053)	17.78%	
Caverage		0.064(0.062)	0.063(0.058)	0.063(0.055)	0.063(0.053)	0.063(0.057)	
	0.1	0.060(0.057)	0.057(0.054)	0.056(0.051)	0.060(0.056)	6.58%	
	0.3	0.071(0.059)	0.074(0.055)	0.074(0.053)	0.075(0.055)	24.62%	
0.3	0.5	0.089(0.064)	0.090(0.054)	0.094(0.054)	0.087(0.050)	38.39%	35.78%
	0.7	0.107(0.061)	0.113(0.053)	0.112(0.052)	0.115(0.051)	51.23%	
	0.9	0.136(0.067)	0.129(0.053)	0.138(0.054)	0.136(0.052)	58.09%	
Caverage		0.092(0.061)	0.093(0.054)	0.095(0.053)	0.095(0.053)	0.094(0.055)	
	0.1	0.064(0.056)	0.063(0.051)	0.064(0.052)	0.062(0.052)	16.33%	
	0.3	0.091(0.056)	0.089(0.053)	0.096(0.052)	0.091(0.049)	42.67%	
0.5	0.5	0.128(0.054)	0.129(0.052)	0.129(0.051)	0.131(0.050)	59.96%	52.21%
	0.7	0.169(0.061)	0.174(0.059)	0.172(0.052)	0.177(0.052)	67.73%	
	0.9	0.220(0.067)	0.229(0.062)	0.220(0.047)	0.226(0.053)	74.33%	
Caverage		0.135(0.059)	0.137(0.055)	0.136(0.051)	0.137(0.052)	0.136(0.054)	
	0.1	0.064(0.053)	0.068(0.052)	0.067(0.051)	0.069(0.047)	24.08%	
	0.3	0.112(0.050)	0.110(0.051)	0.115(0.047)	0.112(0.047)	56.58%	
0.7	0.5	0.176(0.056)	0.174(0.050)	0.173(0.046)	0.174(0.050)	71.21%	62.46%
	0.7	0.240(0.062)	0.250(0.055)	0.249(0.053)	0.253(0.051)	77.62%	
	0.9	0.340(0.066)	0.349(0.063)	0.347(0.057)	0.357(0.053)	82.81%	
Caverage		0.186(0.057)	0.190(0.054)	0.190(0.051)	0.193(0.049)	0.190(0.053)	
	0.1	0.070(0.045)	0.070(0.049)	0.067(0.051)	0.074(0.050)	31.10%	
	0.3	0.133(0.051)	0.137(0.051)	0.137(0.048)	0.133(0.049)	63.19%	
0.9	0.5	0.223(0.051)	0.228(0.053)	0.230(0.051)	0.230(0.047)	77.91%	69.17%
	0.7	0.328(0.051)	0.342(0.053)	0.340(0.054)	0.346(0.049)	84.77%	
	0.9	0.501(0.068)	0.519(0.058)	0.522(0.052)	0.531(0.053)	88.90%	
Caverage		0.251(0.053)	0.259(0.053)	0.259(0.051)	0.263(0.049)	0.258(0.051)	
Overall average		41.64%	45.20%	47.54%	48.66%	45.76%	45.76%

Note: The numbers (and the numbers in bracket) in columns 3 - 6 are "the rejection rate before transformation" (and "the rejection rate after transformation"). 'Raverage' (stands for row average) is the average rejection rate for the same value of α_1 and α_2 but across different values of N ; 'Caverage' (stands for column average) is the average rejection rate for the same value of N and α_2 but across different values of α_1; 'Saverage' (stands for Situation average) is the average rejection rate for the same value of α_2 but across different values of α_1 and N; 'Overall average' in columns 3 - 6 is the overall average for each N for different values of α_1 and α_2 while 'Overall average' in the last two column is the overall average for all the cases in the entire table.

Table F2: Comparison of rejection rates before and after transformation in all cases, error $\sim$ t(5)

α_2	α_1	N=100	N=200	N=400	N=800	Raverage	Saverage
	0.1	0.056(0.057)	0.054(0.056)	0.053(0.052)	0.051(0.050)	-0.23%	
	0.3	0.060(0.060)	0.057(0.056)	0.057(0.053)	0.059(0.051)	5.72%	
0.1	0.5	0.068(0.066)	0.061(0.056)	0.068(0.059)	0.059(0.050)	9.37%	9.37%
	0.7	0.067(0.064)	0.068(0.061)	0.067(0.055)	0.067(0.052)	13.77%	
	0.9	0.068(0.062)	0.072(0.062)	0.077(0.059)	0.075(0.055)	18.24%	
Caverage		0.064(0.062)	0.063(0.058)	0.064(0.055)	0.062(0.052)	0.063(0.057)	
	0.1	0.059(0.059)	0.061(0.054)	0.055(0.051)	0.060(0.052)	8.04%	
	0.3	0.076(0.058)	0.076(0.052)	0.074(0.055)	0.065(0.046)	27.69%	
0.3	0.5	0.089(0.061)	0.090(0.052)	0.095(0.052)	0.093(0.051)	40.86%	37.11%
	0.7	0.112(0.064)	0.116(0.056)	0.113(0.053)	0.115(0.054)	50.19%	
	0.9	0.138(0.068)	0.139(0.059)	0.137(0.049)	0.134(0.050)	58.78%	
Caverage		0.095(0.062)	0.096(0.055)	0.095(0.052)	0.093(0.051)	0.095(0.055)	
	0.1	0.067(0.052)	0.060(0.047)	0.064(0.053)	0.063(0.052)	19.67%	
	0.3	0.090(0.055)	0.096(0.054)	0.088(0.052)	0.091(0.051)	41.73%	
0.5	0.5	0.124(0.055)	0.129(0.052)	0.128(0.054)	0.127(0.051)	58.09%	52.06%
	0.7	0.171(0.065)	0.173(0.057)	0.169(0.049)	0.178(0.054)	67.49%	
	0.9	0.213(0.067)	0.223(0.060)	0.224(0.057)	0.231(0.054)	73.33%	
Caverage		0.133(0.059)	0.136(0.054)	0.134(0.053)	0.138(0.052)	0.135(0.054)	
	0.1	0.067(0.051)	0.067(0.051)	0.069(0.052)	0.065(0.052)	22.95%	
	0.3	0.112(0.056)	0.114(0.050)	0.112(0.050)	0.111(0.050)	54.20%	
0.7	0.5	0.176(0.055)	0.166(0.055)	0.175(0.520)	0.180(0.053)	69.14%	61.24%
	0.7	0.249(0.061)	0.243(0.055)	0.249(0.054)	0.254(0.051)	77.75%	
	0.9	0.334(0.074)	0.349(0.061)	0.345(0.057)	0.343(0.053)	82.15%	
Caverage		0.188(0.060)	0.188(0.054)	0.190(0.053)	0.19(0.052)	0.189(0.055)	
	0.1	0.073(0.052)	0.072(0.049)	0.078(0.051)	0.074(0.048)	33.07%	
	0.3	0.138(0.049)	0.138(0.052)	0.140(0.051)	0.142(0.048)	64.29%	
0.9	0.5	0.225(0.050)	0.228(0.050)	0.216(0.049)	0.229(0.047)	78.17%	69.8%
	0.7	0.345(0.057)	0.338(0.054)	0.354(0.050)	0.349(0.050)	84.78%	
	0.9	0.502(0.072)	0.517(0.059)	0.517(0.053)	0.527(0.049)	88.70%	
Caverage		0.257(0.056)	0.259(0.053)	0.261(0.051)	0.264(0.048)	0.260(0.052)	
Overall average		41.85%	45.79%	47.32%	48.71%	45.92%	45.92%

Note: The numbers (and the numbers in bracket) in columns 3 - 6 are "the rejection rate before transformation" (and "the rejection rate after transformation"). 'Raverage' (stands for row average) is the average rejection rate for the same value of α_1 and α_2 but across different values of N ; 'Caverage' (stands for column average) is the average rejection rate for the same value of N and α_2 but across different values of α_1; 'Saverage' (stands for Situation average) is the average rejection rate for the same value of α_2 but across different values of α_1 and N; 'Overall average' in columns 3 - 6 is the overall average for each N for different values of α_1 and α_2 while 'Overall average' in the last two column is the overall average for all the cases in the entire table.

Table F3: Comparison of rejection rates before and after transformation in all cases, error ∼ t(2)

α_2	α_1	N=100	N=200	N=400	N=800	Raverage	Saverage
	0.1	0.055(0.056)	0.052(0.052)	0.049(0.048)	0.048(0.045)	1.01%	
	0.3	0.055(0.056)	0.053(0.053)	0.051(0.049)	0.051(0.048)	1.93%	
0.1	0.5	0.063(0.062)	0.065(0.062)	0.061(0.053)	0.058(0.049)	8.39%	7.36%
	0.7	0.068(0.066)	0.064(0.058)	0.065(0.056)	0.060(0.049)	10.46%	
	0.9	0.072(0.068)	0.076(0.067)	0.074(0.060)	0.070(0.053)	15.03%	
Caverage		0.062(0.062)	0.062(0.058)	0.060(0.053)	0.057(0.049)	0.06(0.056)	
	0.1	0.060(0.058)	0.054(0.051)	0.056(0.053)	0.051(0.046)	5.34%	
	0.3	0.072(0.058)	0.068(0.055)	0.061(0.048)	0.064(0.046)	22.33%	
0.3	0.5	0.087(0.061)	0.086(0.055)	0.084(0.050)	0.082(0.045)	37.99%	32.92%
	0.7	0.112(0.074)	0.106(0.060)	0.102(0.053)	0.101(0.049)	44.28%	
	0.9	0.139(0.073)	0.129(0.055)	0.128(0.054)	0.126(0.054)	54.66%	
Caverage		0.094(0.065)	0.089(0.055)	0.086(0.052)	0.085(0.048)	0.088(0.055)	
	0.1	0.062(0.052)	0.065(0.054)	0.058(0.048)	0.061(0.049)	17.67%	
	0.3	0.083(0.056)	0.084(0.050)	0.080(0.048)	0.080(0.049)	38.13%	
0.5	0.5	0.119(0.059)	0.115(0.051)	0.107(0.045)	0.110(0.046)	55.52%	50.02%
	0.7	0.161(0.061)	0.156(0.052)	0.151(0.051)	0.151(0.049)	65.9%	
	0.9	0.222(0.071)	0.216(0.059)	0.217(0.054)	0.212(0.051)	72.88%	
Caverage		0.129(0.060)	0.0127(0.053)	0.123(0.049)	0.123(0.049)	0.123(0.053)	
	0.1	0.071(0.055)	0.068(0.047)	0.064(0.047)	0.063(0.046)	26.39%	
	0.3	0.106(0.051)	0.100(0.048)	0.101(0.050)	0.100(0.047)	51.71%	
0.7	0.5	0.161(0.058)	0.159(0.049)	0.153(0.047)	0.155(0.047)	68.06%	61.05%
	0.7	0.234(0.060)	0.234(0.054)	0.230(0.051)	0.226(0.047)	77.09%	
	0.9	0.327(0.072)	0.329(0.059)	0.328(0.054)	0.327(0.051)	82.05%	
Caverage		0.180(0.059)	0.179(0.051)	0.174(0.050)	0.174(0.047)	0.177(0.052)	
	0.1	0.073(0.054)	0.075(0.051)	0.074(0.051)	0.073(0.049)	30.79%	
	0.3	0.134(0.055)	0.129(0.049)	0.122(0.050)	0.128(0.047)	60.78%	
0.9	0.5	0.210(0.053)	0.220(0.050)	0.220(0.050)	0.210(0.047)	76.83%	68.54%
	0.7	0.335(0.056)	0.332(0.047)	0.339(0.046)	0.328(0.048)	85.26%	
	0.9	0.498(0.072)	0.513(0.051)	0.507(0.049)	0.508(0.049)	89.05%	
Caverage		0.250(0.058)	0.254(0.050)	0.252(0.049)	0.250(0.048)	0.251(0.051)	
Overall average		39.42%	43.91%	45.26%	47.33%	43.98%	43.98%

Note: The numbers (and the numbers in bracket) in columns 3 - 6 are "the rejection rate before transformation" (and "the rejection rate after transformation"). 'Raverage' (stands for row average) is the average rejection rate for the same value of α_1 and α_2 but across different values of N ; 'Caverage' (stands for column average) is the average rejection rate for the same value of N and α_2 but across different values of α_1; 'Saverage' (stands for Situation average) is the average rejection rate for the same value of α_2 but across different values of α_1 and N; 'Overall average' in columns 3 - 6 is the overall average for each N for different values of α_1 and α_2 while 'Overall average' in the last two column is the overall average for all the cases in the entire table.

Table F4: Comparison of rejection rates before and after transformation in all cases, error ∼ t(1)

α_2	α_1	N=100	N=200	N=400	N=800	Raverage	Saverage
	0.1	0.046(0.047)	0.039(0.039)	0.028(0.029)	0.025(0.025)	-1.24%	
	0.3	0.048(0.049)	0.042(0.041)	0.035(0.035)	0.031(0.029)	2.13%	
0.1	0.5	0.062(0.062)	0.054(0.053)	0.042(0.041)	0.035(0.033)	2.95%	3.37%
	0.7	0.067(0.066)	0.057(0.056)	0.049(0.045)	0.042(0.038)	4.84%	
	0.9	0.080(0.078)	0.077(0.074)	0.070(0.061)	0.061(0.052)	8.16%	
Caverage		0.061(0.060)	0.054(0.053)	0.045(0.042)	0.039(0.035)	0.050(0.048)	
	0.1	0.054(0.056)	0.044(0.043)	0.035(0.032)	0.030(0.027)	4.23%	
	0.3	0.057(0.053)	0.051(0.042)	0.042(0.032)	0.032(0.024)	18.04%	
0.3	0.5	0.072(0.059)	0.062(0.043)	0.048(0.032)	0.041(0.028)	28.56%	25.17%
	0.7	0.090(0.070)	0.076(0.050)	0.065(0.042)	0.047(0.029)	32.51%	
	0.9	0.120(0.072)	0.114(0.064)	0.097(0.055)	0.079(0.046)	42.54%	
Caverage		0.079(0.062)	0.069(0.048)	0.057(0.039)	0.048(0.031)	0.063(0.045)	
	0.1	0.059(0.054)	0.051(0.044)	0.041(0.036)	0.033(0.029)	11.75%	
	0.3	0.070(0.051)	0.059(0.041)	0.047(0.032)	0.040(0.023)	33.35%	
0.5	0.5	0.086(0.054)	0.073(0.041)	0.064(0.032)	0.053(0.024)	46.40%	42.16%
	0.7	0.125(0.062)	0.107(0.050)	0.083(0.035)	0.069(0.029)	55.32%	
	0.9	0.193(0.071)	0.164(0.059)	0.143(0.051)	0.114(0.042)	63.97%	
Caverage		0.107(0.058)	0.091(0.047)	0.076(0.037)	0.062(0.029)	0.084(0.043)	
	0.1	0.068(0.053)	0.065(0.048)	0.053(0.040)	0.040(0.027)	26.20%	
	0.3	0.092(0.054)	0.078(0.045)	0.067(0.032)	0.053(0.028)	45.80%	
0.7	0.5	0.127(0.058)	0.106(0.041)	0.085(0.029)	0.071(0.026)	61.14%	56.99%
	0.7	0.173(0.055)	0.149(0.042)	0.126(0.032)	0.098(0.023)	72.89%	
	0.9	0.290(0.071)	0.256(0.054)	0.213(0.037)	0.176(0.037)	78.94%	
Caverage		0.150(0.058)	0.131(0.046)	0.109(0.034)	0.088(0.028)	0.119(0.042)	56.99%
	0.1	0.077(0.055)	0.077(0.047)	0.064(0.035)	0.060(0.031)	40.58%	
	0.3	0.121(0.059)	0.111(0.048)	0.095(0.037)	0.082(0.034)	57.11%	
0.9	0.5	0.193(0.051)	0.168(0.044)	0.138(0.033)	0.122(0.030)	74.87%	69.34%
	0.7	0.289(0.051)	0.258(0.039)	0.212(0.031)	0.178(0.027)	84.42%	
	0.9	0.479(0.065)	0.435(0.045)	0.385(0.035)	0.329(0.027)	89.71%	
Caverage		0.232(0.056)	0.21(0.044)	0.179(0.034)	0.154(0.030)	0.194(0.041)	69.34%
Overall average		34.30%	38.65%	41.62%	43.05%	39.41%	39.41%

Note: The numbers (and the numbers in bracket) in columns 3 - 6 are "the rejection rate before transformation" (and "the rejection rate after transformation"). 'Raverage' (stands for row average) is the average rejection rate for the same value of α_1 and α_2 but across different values of N ; 'Caverage' (stands for column average) is the average rejection rate for the same value of N and α_2 but across different values of α_1; 'Saverage' (stands for Situation average) is the average rejection rate for the same value of α_2 but across different values of α_1 and N; 'Overall average' in columns 3 - 6 is the overall average for each N for different values of α_1 and α_2 while 'Overall average' in the last two column is the overall average for all the cases in the entire table.