Could regressing a stationary series on a non-stationary series obtain meaningful outcomes? *

Wing-Keung Wong

Department of Finance and Big Data Research Center, Asia University

Department of Medical Research, China Medical University Hospital, Taiwan

Business, Economic and Public Policy Research Centre,

Hong Kong Shue Yan University

The Economic Growth Centre, Nanyang Technological University

Mu Yue

School of Physical and Mathematical Sciences, Nanyang Technological University

February 8, 2025

Abstract:

We have read many papers in the literature and found that some papers report results of regressing a stationary time series on a non-stationary time series (we call it the IOI1 model). However, very few studies, if there are

*The first author would like to thank Robert B. Miller and Howard E. Thompson for their continuous guidance and encouragement. The first author would also like to thank George Edward Pelham Box and Gregory C. Reinsel for teaching him time series analysis. This paper was done when the first author visited Business, Economic and Public Policy Research Centre and The Economic Growth Centre, Nanyang Technological University in 2024. This research has been supported by Asia University, China Medical University Hospital, Hong Kong Shue Yan University, Nanyang Technological University, Research Grants Council (RGC) of Hong Kong (project numbers 12502814 and 12500915), and the Ministry of Science and Technology (MOST, Project Numbers 106-2410-H-468-002 and 107-2410-H-468-002-MY3), Taiwan.

any, examine the IOI1 model and the robustness of inference in such settings remains an open question. To bridge the gap in the literature, in this paper, we investigate whether regressing a stationary time series, Y_t, on a non-stationary time series, X_t (that is, $Y_t = \alpha + \beta X_t + u_t$) could get any meaningful result. To do so, we first conduct a simulation and find regressing a stationary time series on a non-stationary time series could be spurious.

Thereafter, we develop the estimation and testing theory for the $I0I1$ model and find that the statistics T_N^β for testing $H_0^\beta : \beta = \beta_0$ versus $H_1^\beta : \beta \neq \beta_0$ from the traditional regression model (we call it IOI0 model) does not have any asymptote distribution with $E(T_N^\beta) \to \infty$ and $Var(T_N^\beta) \to \infty$ as $N \to \infty$, and thus, it cannot be used for the $I0I1$ model. We have found other interesting results as shown in our paper. Thus, our paper extends the spurious regression literature to cover a previously unexplored case, thereby contributing to a more comprehensive understanding of time series modeling and inference.

Keywords: Cointegration; stationarity; non-stationarity.

JEL Classification:

1 Introduction

We have read many papers in the literature and found that some papers report results of regressing a stationary time series to a non-stationary time series. For example, Singh et al. (2011) regress stock returns (which is stationary) on GDP (which is not stationary) and found that GDP has positive relationships with stock returns. However, very few studies, if there are any, examine the

IOI1 model and the robustness of inference in such settings remains an open question. To bridge the gap in the literature, in this paper, we investigate whether regressing a stationary time series, Y_t, on a non-stationary time series, X_t, could get any meaningful result. It is well known that a regression applied to nonstationary time series could result in spurious outcomes (Granger and Newbold, 1974).

To investigate whether regressing a stationary time series, Y_t, on a non-stationary time series, X_t, could get any meaningful result, we first call the traditional regression model of regressing a stationary time series on a stationary time series the $I0I0$ model while regressing a stationary time series on a non-stationary time series the $I0I1$ model. We conjecture that the $I0I1$ model could be spurious if one uses the tests from the $I0I0$ model. To check whether our conjecture holds, we conduct a simulation and find that our conjecture holds. Our findings may infer that the tests from the $I0I0$ model do not hold for the $I0I1$ model.

Thereafter, we develop the theory to show that the tests from the $I0I0$ model do not hold for the following $I0I1$ model: $Y_t = \alpha + \beta X_t + u_t$ where u_t is assumed to be iid $N(0, \sigma^2)$, Y_t follows a stationary AR1 model, and X_t is I(1). For this $I0I1$ model, under some assumptions, we obtain the following results: first, $E(\hat{\beta}) = 0$, $E(\widehat{\alpha}) = \mu_Y$, and $E(\hat{Y}_t) = \mu_Y$. Secondly, by using these results, we find that $cov(\overline{X}, S^2) \neq 0$ where $S^2 = \sum_{t=1}^{N} \hat{u}_t^2/(N-2)$ is the estimate of σ^2 and we find that $E(S^2) \to \infty$ and $T \to \infty$. We then find that the t-statistic T_N^{β} for testing $H_0^{\beta} : \beta = \beta_0$ versus $H_1^{\beta} : \beta \neq \beta_0$ does not have any asymptote distribution with $E(T_N^{\beta}) \to \infty$ and $Var(T_N^{\beta}) \to \infty$ as $N \to \infty$, which, in turn, infer that T_N^{β} does not follow t-distribution with $N-2$ degrees of freedom when using it to test the $I0I1$ model, and thus, the test cannot be used to test the $I0I1$ model. In doing so, we extend the spurious regression literature to cover a previously unexplored case, thereby contributing to a more comprehensive understanding of time series modeling and inference.

Section 2 provides background literature on the topic. Section 3 discusses the standard linear regression model and then discuss the model setting for regressing a stationary time series on a non-stationary time series. Section 4 discusses the model setup for the simulation, develops an algorithm for the simulations, and discusses the simulation results by using the algorithm developed in our paper. Section 5 develops the estimation and testing theory for regressing a stationary time series on a non-stationary time series. The last section concludes and suggests future extensions.

2 Literature Review

Time series analysis involves understanding and modeling sequences of data points collected over time at regular intervals. Over the years, various models have been proposed for modeling time series data. Readers may refer to (Tsay, 1989; Nakatani and Teräsvirta, 2009) and others for more information. These models assume that the underlying time series is either stationary or becomes stationary after differencing or other transformations(Brockwell and Davis, 2002). However, when time series data are non-stationary, models such as ARIMA are commonly used, as non-stationary series can present challenges in statistical inference.

One of the key challenges in time series econometrics is the issue of spurious regression, as first noted by Granger and Newbold (1974) who demonstrated that regressing two independent non-stationary time series often results in highly significant coefficients and has high R^2 values, despite no actual causal relationship. This issue was later extended by Phillips (1986), Sun (2004), and others, who provided asymptotic results for spurious regressions. Subsequent works by Ventosa-Santaulária (2009), Marmol (1995), and Kao (1999) further developed the theory and diagnostic tools for detecting spurious regressions, with particular focus on $I(1)$ and fractionally integrated ($I(d)$) processes.

Many studies have highlighted the danger of regressing independent non-stationary series, as it often leads to misleading conclusions (Granger and Newbold, 1974; Phillips, 1986). Agiakloglou (2013) examined spurious regressions in the context of both stationary and non-stationary series, while Kim et al. (2004) explored spurious outcomes when one series has a linear trend, and the other is stationary. The traditional remedy to avoid spurious regression is to either difference the non-stationary series or use cointegration techniques when the series are related but exhibit non-stationarity (Engle and Granger, 1987).

Several studies focus on spurious regression with mixed integration orders. For instance, Pesaran et al. (1999) and Westerlund (2008) studied scenarios where a combination of stationary ($I(0)$) and non-stationary ($I(1)$) series may lead to spurious results. The issue of spurious regression has also been extended to long-memory processes by Tsay and Chung (2000), who considered fractionally integrated ($I(d)$) series, and to scenarios involving seasonal unit roots, as explored by Abeysinghe (1994).

Recently, Cheng, Hui, McAleer, and Wong (2021) conducted simulations and found that under some situations, the regression of two independent and nearly non-stationary series does not have any spurious problem at all while Cheng, Hui, Liu, and Wong (2022) conducted simulations and found that, in some situations, regression is found to be insignificant but actually, they are related. On the other hand, Wong, Cheng, and Yue (2024) conducted simulations and found that regression of stationary time series could be spurious. Moreover, they have proposed a remedial approach to solve the problem.

Despite the vast literature on spurious regression involving non-stationary series, very few studies, if there are any, examine the cases in which one series is stationary, and the other is non-stationary. This gap is particularly relevant because many empirical studies involve mixed-order processes, and the robustness of inference in such settings remains an open question. Our study builds on the foundational work of Granger and Newbold (1974) and seeks to address the overlooked issue of regressing a stationary series on a non-stationary series.

In this paper, we examine whether regressing a stationary time series ($I(0)$) on a non-stationary time series ($I(1)$) can yield meaningful results. We hypothesize that the tests, such as the t-test for β, from the traditional regression model (we call it the $I0I0$ model) designed for stationary series, cannot be used for the $I0I1$ model. To test whether our conjecture holds, we conduct simulation studies and develop an estimation and testing theory for the $I0I1$ model to show that the traditional test statistics cannot be used in this mixed-order setting.

3 The Models

In this paper, we investigate whether regressing a stationary time series, Y_t, on a non-stationary time series, X_t, could get any meaningful result. To do so, in this section, we first discuss the standard linear regression model (we call it the IOI0 model). We then discuss the model setting for regressing a stationary time series on a non-stationary time series (we call it the IOI1 model). In this paper, we call the existing results proposition and theorem if the result is developed in this paper.

3.1 Linear regression model

The earliest form of regression was published by Legendre (1806). When examining the relationship between a quantitative response and a single quantitative explanatory variable, simple linear regression is the most commonly used statistical analysis method. In this paper, we consider the following simple form of regression model:

$$Y_t = \alpha + \beta X_t + u_t, \tag{3.1}$$

where u_t is assumed to be iid $N(0, \sigma^2)$, $t = 1, \cdots, N$ in which N is the sample size, α is the intercept parameter, β is the slope parameter and u_t is a random component denoted as *error term*. Under this standard model setting, one could get the following proposition easily:

Proposition 3.1 *In the above model setting, the least squares estimators $\hat{\beta}$ and $\widehat{\alpha}$*

$$\hat{\beta} = \frac{\sum_{t=1}^{N}(X_t - \bar{X})(Y_t - \bar{Y})}{\sum_{t=1}^{N}(X_t - \bar{X})^2}, \tag{3.2}$$

$$\widehat{\alpha} = \overline{Y} - \widehat{\beta}\overline{X}, \tag{3.3}$$

are the Best Linear Unbiased Estimators (BLUE) of β and α, respectively, such that $E(\widehat{\alpha}) = \alpha$ and $E\left(\widehat{\beta}\right) = \beta$ in the class of both unbiased linear estimators, have minimum variance.

Let the residual $\hat{u}_t = Y_t - \hat{Y}_t$ and

$$S^2 = \frac{\sum_{t=1}^{N} \hat{u}_t^2}{N-2} \tag{3.4}$$

be the estimate of σ^2. Then, we have

$$Var(\hat{\beta}_1) = \frac{\sigma^2}{\sum(X_i - \overline{X})^2}, \quad S^2_{\hat{\beta}_1} = \frac{S^2}{\sum(X_i - \overline{X})^2}, \tag{3.5}$$

$$Var(\hat{\beta}_0) = \sigma^2\left(\frac{\sum {X_i}^2}{N\sum(X_i - \overline{X})^2}\right), \quad S^2_{\hat{\beta}_0} = S^2\left(\frac{\sum {X_i}^2}{N\sum(X_i - \overline{X})^2}\right) \tag{3.6}$$

and one could get the following proposition easily:

Proposition 3.2

(a) $E(\hat{u}) = 0$ *and* $Var(\hat{u}) = \sigma^2$.

(b) $E(S^2) = \sigma^2$;

(c) $s^2_{\hat{\beta}_1} = \dfrac{S^2}{\sum(X_i - \overline{X})^2}$ *is unbiased for* $Var(\hat{\beta}_1)$.

In simple linear regression (3.1), to detect the relationship between two variables, usually, one will test the following hypothesis:

$$H_0^\beta : \beta = \beta_0 \qquad \text{versus} \qquad H_1^\beta : \beta \neq \beta_0 \tag{3.7}$$

If $\beta_0 = 0$ and if the null hypothesis H_0^β in (3.7) is true, then in (3.1) for every value of X_t, the population mean of Y_t always equals to 0. This is to say that Y_t does not depend on the value of X_t, implying that X_t and Y_t have no linear relationship. The alternative is that changes in X_t are associated with changes in Y_t.

Similarly, one may want to test the following hypothesis:

$$H_0^\alpha : \alpha = \alpha_0 \qquad \text{versus} \qquad H_1^\alpha : \alpha \neq \alpha_0 \ . \tag{3.8}$$

If the null hypotheses H_0^β and H_0^α in (3.7) and (3.8), respectively, are true, we have the following proposition:

Proposition 3.3 *In the above model setting, if the null hypotheses* H_0^β *and* H_0^α *in* (3.7) *and* (3.8), *respectively, are true, and if we relax the normality assumption for* u_t, *we have*

$$\hat{\alpha} \xrightarrow{d} N\left(\alpha_0\,,\, \sigma^2\left(\frac{1}{N} + \frac{\overline{X}^2}{\sum(X_i - \overline{X})^2}\right)\right)\,,$$

$$\hat{\beta} \xrightarrow{d} N\left(\beta_0\,,\, \frac{\sigma^2}{\sum(X_i - \overline{X})^2}\right)\,.$$

Proposition 3.4 *In the above model setting,* $\overline{X}$ *and* S^2 *are independent where* $\overline{X}$ *is defined in* (3.2) *and* S^2 *is defined in* (3.4)*.*

To test whether the null hypothesis H_0^β in (3.7) is true, the following T^β test is commonly used:

$$T^\beta = \frac{\hat{\beta} - \beta_0}{SE(\hat{\beta})} \, , \tag{3.9}$$

where $\hat{\beta}$ is the estimate of β in which $\bar{X} = \sum_{t=1}^{N} X_t/N$ and $\bar{Y} = \sum_{t=1}^{N} Y_t/N$ and SE($\hat{\beta}$) is the standard error of the estimate that measures the accuracy of predictions and can be calculated as follows:

$$SE(\hat{\beta}) = \sqrt{\frac{\sum_{t=1}^{N}(\hat{u}_t^2)/(N-2)}{\sum_{t=1}^{N}(X_t - \bar{X})^2}} \tag{3.10}$$

Proposition 3.5 *In the above model setting, if the null hypothesis* H_0^β *in* (3.7) *is true, then* T_N^β *defined in* (3.9) *follows a t-distribution with* $N-2$ *degrees of freedom.*

Similarly, it is well known that if the null hypothesis H_0^α in (3.8) is true, then the following test statistic T_N^α follows a t-distribution with $N-2$ degrees of freedom:

$$T_N^\alpha = \frac{\hat{\alpha} - \alpha_0}{SE(\hat{\alpha})} \, , \tag{3.11}$$

where $\hat{\alpha}$ is the estimate of α in which SE($\hat{\alpha}$) is the standard error of the estimate that measures the accuracy of predictions and can be calculated as follows:

$$SE(\hat{\alpha}) = \sqrt{\frac{\sum_{t=1}^{N}\hat{u}_t^2}{N-2} \times \left(\frac{1}{N} + \frac{\overline{X}^2}{\sum(X_i - \overline{X})^2}\right)} \, . \tag{3.12}$$

Proposition 3.6 *In the above model setting, if the null hypothesis* H_0^α *defined in* (3.8) *is true, then* T_N^α *defined in 3.11 follows a t-distribution with* $N-2$ *degrees of freedom.*

The *goodness-of-fit* or *coefficient of determination R-squared* (R^2):

$$R^2 = \frac{SSR}{SSTO} = 1 - \frac{SSE}{SSTO} , \tag{3.13}$$

is used to measure how good X_t could be used to estimate Y_t 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 $SSE = \sum_{i=1}^{N}(Y_i - \hat{Y}_i)^2$. R-squared is a proportion between 0 and 1. If $R^2 = 1$, the model fits perfectly and in this situation the independent variable X_t could explain 100% of the variation in dependent variable Y_t. On the other hand, if $R^2 = 0$, than X_t accounts for 0% of the variation in Y_t. The model with a larger R^2 value will have more variation in Y_t explained, and therefore is considered a better fitted model. There is no fixed value for an R^2 that assesses all models. The suitable benchmark would depend on the context.

3.2 Regressing stationary series on a nonstationary series

We read many papers in the literature and found that some papers report results of regressing a stationary time series to a non-stationary time series. For example, Singh et al. (2011) find that GDP (which is not stationary) has positive relationships with stock returns (which is stationary). Thus, in this paper, we investigate whether there is any problem to regress a stationary time series to a non-stationary time series by examining the following conjecture:

Conjecture 3.1 *Regression of a stationary Y_t on a non-stationary X_t may not be able to get any meaningful outcome.*

Before we examine the above conjecture, we first examine the following conjecture:

Conjecture 3.2 *Regressing a stationary Y_t on a non-stationary X_t could be spurious if one uses the tests from the standard regression model as shown in*

Section 3.1.

To examine whether regressing a stationary time series to a non-stationary time series could get any meaningful result, we first study whether Conjecture 3.2 holds in the next section, and thereafter, study whether Conjecture 3.1 holds in Section 5.

3.2.1 Model Setup

In this paper, we will use the following linear first order autoregressive AR(1) (Lewis, 1985):

$$Y_t = \phi Y_{t-1} + \varepsilon_t, \quad \varepsilon_t \overset{iid}{\sim} N(0, \sigma_\varepsilon^2) \tag{3.14}$$

with $|\phi| < 1$ in our study because this is the simplest stationary dependent series. The AR(1) model specifies that the output variable depends linearly on its own preceding value. ε_t is uncorrelated with past values of time series. It adds new information to Y_t, and will be denoted as *error* or *innovation* in this paper. We note that when $\phi = 0$, Y_t is equivalent to a white noise series and when $0 < \phi < 1$, Y_t is a stationary time series that exponentially declining to its mean, and when $\phi < 0$, Y_t tends to oscillate between positive and negative values. To examine Conjecture 3.2, we now regress the stationary Y_t defined in (3.14) on a non-stationary X_t such that

$$X_t = X_{t-1} + e_t, \quad e_t \overset{iid}{\sim} N(0, \sigma_e^2) \; . \tag{3.15}$$

We note that one could use more complicated distributions for both ε_t and e_t. In this paper, we only consider a simple case in which both ε_t and e_t are normally distributed and ε_i and e_j are independent for $i \neq j$, and when i = j we have $\text{Cor}(\varepsilon_i, e_i) = \rho$. Now, we impose the following conditions to X_t and Y_t:

$$X_0 = 0, Y_0 = \mu \; . \tag{3.16}$$

4 Simulation

To study whether Conjecture 3.2 holds, we first set up a model for the simulation as shown in this section. We then develop an algorithm for our simulations and discuss the simulation results by using the algorithm developed in our paper.

4.1 Model Setup

We first set up a model for the simulation. In this paper, we consider a stationary series to be weakly stationary or covariance stationary. One could easily extend our result to include strictly stationary series.

We will use (3.14) to define Y_t and use (3.15) to define X_t. We note that one could use more complicated distributions for both ε_t and e_t. In this paper, we only consider a simple case in which both ε_t and $e_t \sim$ N(0,1). However, in order to simulate X_t and Y_t properly, without loss of generality we will consider different factors that may affect properties of time series carefully.

First, we control the lengths of times series. Because longer series will include more information than shorter ones, in this paper, we simulate the time series with the following four different lengths in our study:

(i) N=100, (ii) N=500, (iii) N=1000.

After deciding the lengths of both X_t and Y_t, we now consider the different values of ϕ:

(a) $\phi = 0$, (b) $\phi = 0.1$, (c) $\phi = 0.3$, (d) $\phi = 0.5$, (e) $\phi = 0.7$, (f) $\phi = 0.9$.

For simplicity, we only consider positive coefficient of $|\phi|$ in this paper.

With 3 different time series lengths, and the above 6 combinations of ϕ values, there are in total 18 subcases of simulation in our study.

4.2 Algorithm

We turn to develop an algorithm for our simulations. Since X_t and Y_t are generated independently, they are not related and thus one may expect the slope coefficient estimator $\hat{\beta}$ is not significantly different from zero. For each subcase (different error distributions, different time series lengths, different combinations of α_1 and α_2) described in Section 3.2.1, we consider the following algorithm:

1. Simulate 10000 pairs of X_t and Y_t defined in (3.14) with coefficients described in Section 3.2.1.

2. Fit model (3.1) for each pair of simulated X_t and Y_t. Therefore, in each subcase, we will obtain 10000 $\hat{\beta}$'s and 10000 corresponding p-values.

3. Use the T test defined in (3.9) to test whether the null hypothesis H_0 in (3.7) 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 subcase. This proportion is denoted as *rejection rate* in this report.

The above algorithm helps to examine whether the t statistic in (3.9) for the model in (3.1) follow a Student t-distribution. If X_t and Y_t are unrelated, the true null hypothesis that all β coefficients are zero should be rejected 5% of the time at the significance level of 0.05. If t-test is perfect, that is, $\hat{\beta}$'s follow student t-distribution, the rejection rate will be exactly 0.05. If rejection rate is significantly greater than 0.05, it indicates that we could not use results from t-test to decide if two variables are truly related because the null hypothesis of no relationship is being rejected too easily.

4.3 Simulation Results

We turn to discuss the simulation results by using the algorithm developed in our paper. Rejection rates from the model are presented in Table 4.1. In case (a), when $\phi = 0$, Y_t is simply random error and thus non-stationary, regressing it on a stationary X_t results in a rejection rate of about 5%, indicating that the t-test can yield valid results for OLS coefficients in this specific scenario. However, in all other cases, more than 6% of the independent variables are significant at the 0.05 level of significance. when the sample size N is increasing, the rejection rate are quite stable, meaning that the rejection rate will not affected by sample size N a lot. When $|\phi|$ is increasing, from stronger stationary to weak or non-stationary time series Y, the rejection rate in increasing. In particular, when ϕ = 0.9, the average rejection rates across different sample size could reach around 60%, suggesting that out of 10,000 regressions, the independent variable X_t is significant 60% of the time. However, since we simulated independent series, we are sure that X_t is not related to Y_t at all and $\beta = 0$. We can conclude that when regressing a stationary Y_t on a non-stationary X_t, the t-test may not provide a valid analysis for OLS coefficients.

Table 4.1: Rejection Rate in the simulation for various ϕ and N.

Case	Coefficients	N=100	N=500	N=1000	Average
(a)	$\phi = 0$	0.0502	0.0466	0.0519	0.0502
(b)	$\phi = 0.1$	0.0718	0.0708	0.0780	0.0718
(c)	$\phi = 0.3$	0.1361	0.1424	0.1521	0.1361
(d)	$\phi = 0.5$	0.2378	0.2528	0.2532	0.2378
(e)	$\phi = 0.7$	0.3819	0.4067	0.4078	0.3819
(f)	$\phi = 0.9$	0.5955	0.6447	0.6449	0.5955

5 The Theory

In this section, we develop the estimation and testing theory for regressing a stationary time series on a non-stationary time series (we call it the IOI1 model). Before we develop the theory, we first state the following proposition:

Proposition 5.1 *Under the model setting stated in Section 3.2.1, we have*

$$X_t = \sum_{i=1}^{t} e_i \ . \tag{5.1}$$

We make the following assumption in this paper:

Assumption 5.1 *Under the model setting stated in Section 3.2.1, in this paper, we assume* $\sum_{t=1}^N \sum_{i=1}^t e_i Y_t$ *and* $\sum_{i=t}^N \sum_{i=1}^t e_i^2$ *are independent.*

Theorem 5.1 *Under the model setting stated in Section 3.2.1, if* Y_t *and* X_t *are defined by* (3.14) *and* (3.15)*, respectively, and Assumption 5.1 holds, then* $E(\hat{\beta}) = 0$ *and* $E(\widehat{\alpha}) = \mu_Y$ *as* $T \to \infty$ *where* $\hat{\beta}$ *and* $\hat{\alpha}$ *are defined by* (3.2) *and* (3.3)*, respectively.*

Proof. From (3.2), we have

$$\widehat{\beta} = \frac{\sum_{t=1}^{N} \left(X_t - \overline{X}\right)\left(Y_t - \overline{Y}\right)}{\sum_{t=1}^{N} \left(X_t - \overline{X}\right)^2} \ .$$

Now, without loss of generation, we assume $\overline{X} = 0$ and $\overline{Y} = \mu_Y$. Thus, we have

$$\begin{aligned} \widehat{\beta} &= \frac{\sum_{t=1}^{N} X_t \left(Y_t - \mu_Y\right)}{\sum_{t=1}^{N} X_t^2} \\ &= \frac{\sum_{t=1}^N \sum_{i=1}^t e_i Y_t}{\sum_{t=1}^N \sum_{i=1}^t e_i^2} \ . \end{aligned}$$

$$
\begin{aligned}
\Rightarrow E(\widehat{\beta}) &= E\left(\frac{\sum\limits_{t=1}^{N}\sum_{i=1}^{t} e_i Y_t}{\sum\limits_{t=1}^{N}(\sum_{i=1}^{t} e_i)^2}\right) \\
&= \frac{E\left(\sum\limits_{t=1}^{N}\sum_{i=1}^{t}(\phi^t \mu e_i + \sum_{j=0}^{t}\phi^j \varepsilon_{t-j} e_i)\right)}{E\left(\sum\limits_{t=1}^{N}(\sum_{i=1}^{t} e_i)^2\right)} \\
&= \frac{\sum\limits_{t=1}^{N}\sum_{i=1}^{t}\sum_{j=0}^{t}\phi^j E\left(\varepsilon_{t-j} e_i\right)}{\sum\limits_{t=1}^{N} E\left((\sum_{i=1}^{t} e_i)^2\right)} \\
&= \frac{\sum\limits_{t=1}^{N}\sum_{i=1}^{t}\phi^{t-i} Cov(\varepsilon_i, e_i)}{\sum\limits_{t=1}^{N} E\left((\sum_{i=1}^{t} e_i)^2\right)} \\
&= \frac{\sum\limits_{t=1}^{N}\rho\sigma_e\sigma_\varepsilon\sum\limits_{i=0}^{t-1}\phi^i}{\sum\limits_{t=1}^{N} t\sigma_e^2} \\
&= \frac{\rho\sigma_e\sigma_\varepsilon\sum\limits_{t=1}^{N}\dfrac{1-\phi^t}{1-\phi}}{\frac{N(N+1)}{2}\sigma_e^2} \\
&= \frac{\rho\sigma_\varepsilon}{(1-\phi)\sigma_e}\frac{\sum\limits_{t=1}^{N} 1-\phi^t}{\frac{N(N+1)}{2}} \\
&= \frac{\rho\sigma_\varepsilon}{(1-\phi)\sigma_e}\frac{N - \dfrac{\phi(1-\phi^T)}{1-\phi}}{\frac{N(N+1)}{2}} \\
&\to \frac{\rho\sigma_\varepsilon}{(1-\phi)\sigma_e} \times 0 = 0 \text{ when } T \to \infty \ .
\end{aligned}
\tag{5.2}
$$

$E(\widehat{\alpha}) = \mu_Y$ is immediate from taking $\overline{X} = 0$.

■

Theorem 5.2 *Under the model setting stated in Section 3.2.1, if Y_t and X_t are defined by* (3.14) *and* (3.15)*, respectively, and Assumption 5.1 holds, then* $E(\hat{Y}_t) = \mu_Y$.

Proof.

$$
\begin{aligned}
E(\hat{Y}_t) &= E(\hat{\alpha}) + E(\hat{\beta})E(X_t) \\
&= \mu_Y + 0 \times 0 \\
&= \mu_Y \ .
\end{aligned}
$$

■

Theorem 5.3 *Under the model setting stated in Section 3.2.1, if Y_t and X_t are defined by* (3.14) *and* (3.15)*, respectively, and Assumption 5.1 holds, then* $E(S^2) \to \infty$ *and* $N \to \infty$ *where* S^2 *is defined by* (3.4).

Proof. Let the residual $\hat{u}_t = Y_t - \hat{Y}_t$, from (3.4), the estimate of σ^2 is

$$
\begin{aligned}
S^2 &= \frac{\sum_{t=1}^{N} \hat{u}_t^2}{N-2} \\
&= \frac{\sum_{t=1}^{N} \left(Y_t - \hat{Y}_t\right)^2}{N-2} \ .
\end{aligned}
$$

Since $N \to \infty$, without loss of generality, we let $\hat{Y}_t = \mu_Y$. Thus, we have

$$
\begin{aligned}
S^2 &= \frac{\sum_{t=1}^{N} \left(Y_t - \hat{Y}_t\right)^2}{N-2} \\
&= \frac{\sum_{t=1}^{N} \left(Y_t - \mu_Y\right)^2}{N-2} \\
&= \frac{\sum_{t=1}^{N} \left(\alpha + \beta X_t + u_t - \mu_Y\right)^2}{N-2} \\
&= \frac{\sum_{t=1}^{N} \left(\alpha + \beta \sum_{i=1}^{t} e_i + u_t - \mu_Y\right)^2}{N-2} \ .
\end{aligned}
$$

Hence,

$$
\begin{aligned}
E(S^2) &= E\left(\frac{\sum_{t=1}^{N}\left(\alpha+\beta\sum_{i=1}^{t}e_i+u_t-\mu_Y\right)^2}{N-2}\right)\\
&= E\left(\frac{\sum_{t=1}^{N}\left(\beta\sum_{i=1}^{t}e_i+u_t\right)^2}{N-2}\right)\\
&= \frac{\beta^2\sum_{t=1}^{N}E((\sum_{i=1}^{t}e_i)^2)+E(\sum_{t=1}^{N}u_t^2)}{N-2}\\
&= \frac{\beta^2\sum_{t=1}^{N}t\sigma_e^2+(\sum_{t=1}^{N}\sigma^2)}{N-2}\\
&= \frac{\beta^2\frac{N(N+1)}{2}\sigma_e^2+N\sigma^2}{N-2}\to\infty \text{ when } N\to\infty\ .
\end{aligned}
$$

■

Theorem 5.4 *Under the model setting stated in Section 3.2.1, if Y_t and X_t are defined by* (3.14) *and* (3.15)*, respectively, then $cov(\overline{X}, S^2) \neq 0$ where $\overline{X}$ is defined in* (3.2) *and S^2 is defined in* (3.4)*.*

Proof.

$$
\begin{aligned}
cov(\overline{X}, S^2) &= E(\overline{X}S^2) - E(\overline{X})E(S^2)\\
&= E\left(\overline{X}\cdot\frac{\sum_{t=1}^{N}(Y_t-\hat{Y}_t)^2}{N-2}\right) - \mu_X E(S^2)\ .
\end{aligned}
$$

Without loss of generality, we let $\hat{Y}_t = \mu_Y$ and let $Y_t = \alpha + \beta X_t + u_t$. Then,

$$
\begin{aligned}
cov(\overline{X}, S^2) &= E\left(\frac{\sum_{t=1}^{N}X_t}{T}\cdot\frac{\sum_{t=1}^{N}(\alpha+\beta X_t+u_t-\mu_Y)^2}{N-2}\right) - 0\times E(S^2)\\
&= \frac{E\left((\sum_{t=1}^{N}X_t)\cdot\sum_{t=1}^{N}(\alpha+\beta X_t+u_t-\mu_Y)^2\right)}{N(N-2)}\ .
\end{aligned}
$$

From the above, obviously, we can conclude that $cov(\overline{X}, S^2) \neq 0$. ■

Theorem 5.5 *Under the model setting stated in Section 3.2.1, if Y_t and X_t are defined by* (3.14) *and* (3.15)*, respectively, and Assumption 5.1 holds, then T_N^β defined in* (3.9) *does not have any asymptote distribution with $E(T_N^\beta) \to \infty$ and $Var(T_N^\beta) \to \infty$ as $N \to \infty$.*

Proof.

Without loss of generality, we assume $\beta_0 = 0$ so that Equation (3.9) becomes

$$T_N^\beta = \frac{\hat{\beta}}{SE(\hat{\beta})} \tag{5.3}$$

$$= \frac{\frac{\sum_{t=1}^N \sum_{i=1}^t e_i Y_t}{\sum_{t=1}^N \sum_{i=1}^t e_i^2}}{\sqrt{\frac{\sum_{t=1}^N (\hat{u}_t^2)/(N-2)}{\sum_{t=1}^N (X_t - \bar{X})^2}}} \tag{5.4}$$

$$= \frac{\left(\sum_{t=1}^N \sum_{i=1}^t e_i Y_t\right) \sqrt{\sum_{t=1}^N (X_t - \bar{X})^2}}{\left(\sum_{t=1}^N \sum_{i=1}^t e_i^2\right) \sqrt{\sum_{t=1}^N (\hat{u}_t^2)/(N-2)}} \tag{5.5}$$

$$= \frac{\left(\sum_{t=1}^N \sum_{i=1}^t e_i Y_t\right) \sqrt{\sum_{t=1}^N \sum_{i=1}^t e_i^2}}{\left(\sum_{t=1}^N \sum_{i=1}^t e_i^2\right) \sqrt{\sum_{t=1}^N (\hat{u}_t^2)/(N-2)}} \tag{5.6}$$

$$= \frac{\left(\sum_{t=1}^N \sum_{i=1}^t e_i Y_t\right) \sqrt{\sum_{t=1}^N \sum_{i=1}^t e_i^2}}{\left(\sum_{t=1}^N \sum_{i=1}^t e_i^2\right) \sqrt{\sum_{t=1}^N (Y_t - \hat{Y}_t)^2/(N-2)}} \tag{5.7}$$

$$= \frac{\left(\sum_{t=1}^N \sum_{i=1}^t e_i Y_t\right) \sqrt{\sum_{t=1}^N \sum_{i=1}^t e_i^2}}{\left(\sum_{t=1}^N \sum_{i=1}^t e_i^2\right) \sqrt{\sum_{t=1}^N (\alpha + \beta X_t + u_t - \mu_Y)^2/(N-2)}} \tag{5.8}$$

$$= \frac{(N-2) \sum_{t=1}^N \sum_{i=1}^t e_i Y_t}{\sqrt{\left(\sum_{t=1}^N \sum_{i=1}^t e_i^2\right) \left(\sum_{t=1}^N (\alpha + \beta X_t + u_t - \mu_Y)^2\right)}}\,. \tag{5.9}$$

(5.10)

We let $K = \alpha - \mu_Y$, $Y_t = \alpha + \beta X_t + u_t$, and $X_t = \sum_{i=1}^{t} e_i$, we have

$$T_N^{\beta} = \frac{(N-2)\sum_{t=1}^{N}\sum_{i=1}^{t} e_i(\alpha + \beta\sum_{i=1}^{t} e_i + u_t)}{\sqrt{\left(\sum_{t=1}^{N}\sum_{i=1}^{t} e_i^2\right)\left(\sum_{t=1}^{N}\left(K + \beta\sum_{i=1}^{t} e_i + u_t\right)^2\right)}} + \delta \tag{5.11}$$

$$\tag{5.12}$$

We have

$$E(T_N^\beta) = E\left(\frac{(N-2)\sum_{t=1}^N \sum_{i=1}^t e_i(\alpha + \beta \sum_{i=1}^t e_i + u_t)}{\sqrt{\left(\sum_{t=1}^N \sum_{i=1}^t e_i^2\right)\left(\sum_{t=1}^N (K + \beta \sum_{i=1}^t e_i + u_t)^2\right)}}\right) \tag{5.13}$$

$$= (N-2)E\left(\frac{\sum_{t=1}^N \sum_{i=1}^t e_i(\alpha + \beta \sum_{i=1}^t e_i + u_t)}{\sqrt{\left(\sum_{t=1}^N \sum_{i=1}^t e_i^2\right)\left(\sum_{t=1}^N (K + \beta \sum_{i=1}^t e_i + u_t)^2\right)}}\right) \tag{5.14}$$

$$= (N-2)E\left(\frac{\alpha(\sum_{t=1}^N \sum_{i=1}^t e_i) + \beta \sum_{t=1}^N (\sum_{i=1}^t e_i)^2 + \sum_{t=1}^N u_t \sum_{i=1}^t e_i}{\sqrt{\left(\sum_{t=1}^N \sum_{i=1}^t e_i^2\right)\left(\sum_{t=1}^N (K + \beta \sum_{i=1}^t e_i + u_t)^2\right)}}\right) \tag{5.15}$$

$$= (N-2)\frac{E\left(\alpha(\sum_{t=1}^N \sum_{i=1}^t e_i) + \beta \sum_{t=1}^N (\sum_{i=1}^t e_i)^2 + \sum_{t=1}^N u_t \sum_{i=1}^t e_i\right)}{E\left(\sqrt{\left(\sum_{t=1}^N \sum_{i=1}^t e_i^2\right)\left(\sum_{t=1}^N (K + \beta \sum_{i=1}^t e_i + u_t)^2\right)}\right)} + \delta \tag{5.16}$$

$$= (N-2)\frac{\alpha E\left(\sum_{t=1}^N \sum_{i=1}^t e_i)\right) + \beta E\left(\sum_{t=1}^N (\sum_{i=1}^t e_i)^2\right) + E\left(\sum_{t=1}^N u_t \sum_{i=1}^t e_i\right)}{E\left(\sqrt{\left(\sum_{t=1}^N \sum_{i=1}^t e_i^2\right)\left(\sum_{t=1}^N (K + \beta \sum_{i=1}^t e_i + u_t)^2\right)}\right)} + \delta \tag{5.17}$$

$$= (N-2)\frac{0 + \beta E\left(\sum_{t=1}^N (\sum_{i=1}^t e_i)^2\right) + 0}{E\left(\sqrt{\left(\sum_{t=1}^N \sum_{i=1}^t e_i^2\right)\left(\sum_{t=1}^N (K + \beta \sum_{i=1}^t e_i + u_t)^2\right)}\right)} + \delta \tag{5.18}$$

$$= (N-2)\frac{\beta \sigma_e^2 N(N+1)/2}{E\left(\sqrt{\left(\sum_{t=1}^N \sum_{i=1}^t e_i^2\right)\left(\sum_{t=1}^N (K + \beta \sum_{i=1}^t e_i + u_t)^2\right)}\right)} + \delta \tag{5.19}$$

$$= \frac{O(N^3)}{\ldots} + \delta\ . \tag{5.20}$$

We now work on the denominator, D, of the above expression as shown in the following

$$D = E\left(\sqrt{\left(\sum_{t=1}^{N}\sum_{i=1}^{t} e_i^2\right)\left(\sum_{t=1}^{N}(K + \beta\sum_{i=1}^{t} e_i + u_t)^2\right)}\right) \tag{5.21}$$

$$\approx \sqrt{E\left(\sum_{t=1}^{N}\sum_{i=1}^{t} e_i^2\right) \cdot E\left(\sum_{t=1}^{N}\left(K + \beta\sum_{i=1}^{t} e_i + u_t\right)^2\right)} \tag{5.22}$$

$$\approx \sqrt{\frac{\sigma_e^2 N(N+1)}{2} \cdot \left(TK^2 + \beta^2\sigma_e^2\frac{N(N+1)}{2} + T\sigma^2\right)} \tag{5.23}$$

$$= O(N^2)\ . \tag{5.24}$$

Thus,

$$E(T_N^\beta) = \frac{O(N^3)}{O(N^2)} = O(N) \to \infty \text{ as } N \to \infty\ . \tag{5.25}$$

On the other hand,

$$Var\left(T_N^\beta\right) = E\left[\left(T_N^\beta - E(T_N^\beta)\right)^2\right] \to \infty$$

because $E(T_N^\beta) \to \infty$ as $N \to \infty$. $Var\left(T_N^\beta\right) < \infty$ only if $T_N^\beta = E(T_N^\beta) + k$ but this is impossible. Thus, $Var(T_N^\beta) \to \infty$ as $N \to \infty$.

■

Property 5.1 *Under the model setting stated in Section 3.2.1, if Y_t and X_t are defined by* (3.14) *and* (3.15), *respectively, and Assumption 5.1 holds, then T_N^β defined in* (3.9) *does not follow t-distribution with $N-2$ degrees of freedom.*

Theorem 5.6 *If Y_t and X_t are defined by* (3.14) *and* (3.15), *respectively, and Assumption 5.1 holds, then T_N^α defined in 3.11 does not have any asymptote distribution with $E(T_N^\alpha) \to \infty$ and $Var(T_N^\alpha) \to \infty$ as $N \to \infty$.*

The problem in real data analysis, analysts could not be able to know the equation of both Y_t and X_t are from equation (3.14) and thus they cannot get the distribution of $\hat{\beta}$. In this situation, we recommend to use the remedy approach we discuss in next section.

Property 5.2 *If Y_t and X_t are defined by* (3.14) *and* (3.15)*, respectively, and Assumption 5.1 holds, then T_N^α defined in 3.11 does not follow a t-distribution with $N-2$ degrees of freedom.*

6 Conclusion and Future Study

We have read many papers in the literature and found that some papers report results of regressing a stationary time series on a non-stationary time series (we call it the IOI1 model). For example, Singh et al. (2011) regress stock returns (which is stationary) on GDP (which is not stationary) and found that GDP has positive relationships with stock returns. However, very few studies, if there are any, examine the IOI1 model and the robustness of inference in such settings remains an open question. To bridge the gap in the literature, in this paper, we investigate whether regressing a stationary time series, Y_t, on a non-stationary time series, X_t, could get any meaningful result. It is well known that a regression applied to nonstationary time series could result in spurious outcomes (Granger and Newbold, 1974).

To investigate whether regressing a stationary time series, Y_t, on a non-stationary time series, X_t, could get any meaningful result, we first call the traditional regression model of regressing a stationary time series on a stationary time series the $I0I0$ model while regressing a stationary time series on a non-stationary time series the $I0I1$ model. We conjecture that the $I0I1$ model could be spurious if one uses the tests from the $I0I0$ model. To check whether our conjecture holds, we conduct a simulation and find that our conjecture holds. Our findings may infer that the tests from the $I0I0$ model do not hold for the $I0I1$ model.

Thereafter, we develop the theory to show that the tests from the $I0I0$ model do not hold for the following $I0I1$ model: $Y_t = \alpha + \beta X_t + u_t$ where u_t is assumed to be iid $N(0, \sigma^2)$, Y_t follows a stationary AR1 model, and X_t is I(1). For this $I0I1$ model, under some assumptions, we obtain the following results: first, $E(\hat{\beta}) = 0$, $E(\widehat{\alpha}) = \mu_Y$, and $E(\hat{Y}_t) = \mu_Y$. Secondly, by using these results, we find that $cov(\overline{X}, S^2) \neq 0$ where $S^2 = \sum_{t=1}^{N} \hat{u}_t^2/(N-2)$ is the estimate of σ^2 and we find that $E(S^2) \to \infty$ and $T \to \infty$. We then find that the t-statistic T_N^{β} for testing $H_0^{\beta}: \beta = \beta_0$ versus $H_1^{\beta}: \beta \neq \beta_0$ does not have any asymptote distribution with $E(T_N^{\beta}) \to \infty$ and $Var(T_N^{\beta}) \to \infty$ as $N \to \infty$, which, in turn, infer that T_N^{β} does not follow t-distribution with $N-2$ degrees of freedom when using it to test the $I0I1$ model, and thus, the test cannot be used to test the $I0I1$ model. In doing so, we extend the spurious regression literature to cover a previously unexplored case, thereby contributing to a more comprehensive understanding of time series modeling and inference.

Further research includes developing the theory, including the estimation and testing theories, for the $I0I1$ model.

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.

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.

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

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.

Legendre, A. M. (1806), 'Nouvelles méthodes pour la détermination des orbites des cometes (1805)', *New Methods for the Determination of Comet Orbits* .

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

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.

Singh, T., Mehta, S. and Varsha, M. (2011), 'Macroeconomic factors and stock returns: Evidence from taiwan', *Journal of economics and international finance* **3**(4), 217.

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

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.

Ventosa-Santaulária, 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.