Could regressing a stationary series on a non-stationary series obtain meaningful outcomes – A remedy *

Wing-Keung Wong

Department of Finance, Quantum AI Research Center,

Fintech & Blockchain Research Center,

and Big Data Research Center, Asia University

Department of Medical Research, China Medical University Hospital

Business, Economic and Public Policy Research Centre,

Hong Kong Shue Yan University

The Economic Growth Centre, Nanyang Technological University

Minh Tam Pham

Department of Finance, College of Management, Asia University

Mu Yue

School of Physical and Mathematical Sciences, Nanyang Technological University

Abstract:

In this paper, we investigate whether 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 from the standard regression model $Y_t = \alpha + \beta X_t + u_t$ where u_t is assumed to be iid $N(0, \sigma^2)$ could be used for regression with autoregressive noise. To do so, we set $Y_t = Y_{1,t} + Y_{2,t}$ with $Y_{2,t} = \phi Y_{2,t-1} + e_t$ in which $e_t \overset{iid}{\sim} (0, \sigma_e^2)$ so that the regression contains autoregressive noise and we use the statistics T_N^β for testing $H_0^\beta : \beta = \beta_0$ versus $H_1^\beta : \beta \neq \beta_0$ when the actual $\beta > 0$, for example, $\beta = 0.1$. In our simulation, we found that the average rejection rate is less than the level of significance for any sample size N smaller than 100. However, for large N, say, $N = 1000$, the test confirms that the model is significant. Our findings confirm that the test from the standard regression model could make a significant regression with autoregressive noise become insignificant for small sample sizes, but not for very large sample sizes.

Keywords: Stationarity, autoregression, regression, time series analysis

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

*The first author would like to thank Robert B. Miller and Howard E. Thompson for their continuous guidance and encouragement. We would also like to thank Shing Fung Chan for his assistance in our paper. The first author would also like to thank George Edward Pelham Box and Gregory C. Reinsel for teaching him time series analysis. This project was done when I visited Hong Kong Shue Yan University and Nanyang Technological University in 2024. This research is partially supported by grants from 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.

1 Introduction

Based on the foundational work of Granger and Newbold (1974) to address regressing a stationary series, Y_t, on a non-stationary series, X_t, (we call it the I0I1 model), many papers in the literature use the I0I1 model. For example, Singh et al. (2011) regressed stock returns (which are stationary) on GDP (which is not stationary). However, very few studies, if there are any, have examined whether the I0I1 model has problems. To distinguish the I0I1 model, we call the traditional regression model of regressing a stationary time series on a stationary time series the $I0I0$ model. To test whether using the statistics T_N^β from the $I0I0$ model can be used for testing $H_0^\beta : \beta = \beta_0$ versus $H_1^\beta : \beta \neq \beta_0$ in the $I0I1$ model, Wong and Yue (2024) conducted simulations and found that the I0I1 model could be spurious. They also show that the statistics T_N^β from the I0I0 model cannot be used for the $I0I1$ model. To bridge the gap in the literature, in this paper, we introduced three remedies to overcome the limitations of the statistics.

Our first proposed remedy (Remedy 1) is to recommend academics and practitioners regress Y_t on $X_t' = X_t - X_{t-1}$. We then conduct simulations to examine the performance of Remedy 1. We find that using Remedy 1 could correct the spurious problem, but it results in getting a much smaller size. Thus, we conclude that Remedy 1 is not a good remedy.

We then develop Remedy 2 and recommend it to academics and practitioners to regress $Y_t' = (1 - \hat{\phi}_Y B)Y_t$ on $X_t' = X_t - X_{t-1}$. We then conduct simulations to examine the performance of Remedy 2 and find that Remedy 2 not only can correct the spurious problem, but also can keep the size close to the theoretical benchmark. Thus, we conclude that Remedy 2 is a very good remedy.

Last, we develop Remedy 3 and recommend it to academics and practitioners to regress $Y_t'' = \frac{Y_t}{1-\phi_Y B}$ on $X_t'' = \frac{X_t}{1-\phi_X B}$, conduct simulations to examine the performance of Remedy 3, and find that, similar to Remedy 2, Remedy 3 can correct the spurious problem and keep the size close to the theoretical benchmark. Thus, we conclude that Remedy 3 is also a very good remedy. To examine between Remedy 2 and Remedy 3, which one is better, we conducted simulation and found that the overall average of the rejection rate by using Remedy 2 is 0.0484, while the overall average of the rejection rate by using Remedy 3 is 0.0496, which is closer to the theoretical benchmark. Thus, we conclude that Remedy 3 is slightly better according to the results of our simulation.

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 and develops an algorithm for the simulations while Section 5 discusses the simulation results by using the algorithm developed in our paper. 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. Several models, see, for example, (Tsay, 1989; Nakatani and Teräsvirta, 2009) have been developed to use time series to analyze stationary data. These models can be used for nonstationary data if the data becomes stationary after differencing or other transformations(Brockwell and Davis, 2002). In addition, ARIMA models can be used for nonstationary series.

One of the most important issues in time series analysis is how to test a spurious model. For example, Granger and Newbold (1974) has shown that regressing two independent non-stationary time series could get highly significant coefficients and high R^2 values, though the series are independent. Phillips (1986), Sun (2004), and others have developed asymptotic results for spurious regressions and Ventosa-Santaulária (2009), Marmol (1995), and Kao (1999) further extended the theory by providing diagnostic tools for detecting spurious regressions for $I(1)$ and fractionally integrated ($I(d)$) processes with $0 < d < 1$.

(Granger and Newbold, 1974; Phillips, 1986) and others have concluded that regressing independent non-stationary series could lead to incorrect conclusions. Other studies in this area include Agiakloglou (2013) and Kim et al. (2004). (Engle and Granger, 1987) proposed either differencing the non-stationary series or using cointegration techniques to avoid spurious regression. On the other hand, Pesaran et al. (1999) and Westerlund (2008) found that a combination of stationary ($I(0)$) and non-stationary ($I(1)$) series could lead to spurious results. Furthermore, Tsay and Chung (2000) extended the theory by considering fractionally integrated ($I(d)$) series to model long-memory processes, while Abeysinghe (1994) extended the theory to involve seasonal unit roots.

Recently, conducting simulations, Cheng et al. (2021) found that the regression of two independent and nearly non-stationary series may not have any spurious problem in some situations, Cheng et al. (2022) found that insignificant regression could be significantly related, Wong et al. (2024) found that regression of stationary time series could be spurious, Wong and Yue (2024) found that regressing a stationary series on a non-stationary series may not get any meaningful outcomes, and Wong and Pham (2024a) found that the correlation of a stationary series with a non-stationary series may not get any meaningful outcomes.

Based on the foundational work of Granger and Newbold (1974) to address the I0I1 model, many papers in the literature use the I0I1 model. For example, Singh et al. (2011) regressed stock returns on GDP. Recently, Wong and Yue (2024) conducted simulations and found that the I0I1 model could be spurious, and the statistics T_N^β from the I0I0 model cannot be used for the $I0I1$ model. To extend their work, in this paper, we introduced three remedies to overcome the limitations of the statistics.

3 The Models

Wong and Yue (2024) conclude that regressing a stationary time series, Y_t, on a non-stationary time series, X_t, could yield any meaningful result. To bridge the gap in the literature, in this paper, we provide a remedy approach and show that by using the remedy approach, the test can be used to test whether Y_t and X_t are significantly related. To do so, we first discuss the standard linear regression model (we call it the I0I0 model) in this section.

3.1 Linear regression model

In this paper, we follow Wong and Yue (2024) to consider the following simple linear regression model and call it I0I0 model:

$$Y_t = \alpha + \beta X_t + u_t \ , \quad t = 1, \cdots, N; \tag{3.1}$$

in which u_t is the *error term* assumed to be iid $N(0, \sigma^2)$, N is the sample size, and β is slope parameter with the following estimate $\hat{\beta}$:

$$\begin{aligned} \hat{\beta} &= \frac{\sum_{t=1}^{N}(X_t - \bar{X})(Y_t - \bar{Y})}{\sum_{t=1}^{N}(X_t - \bar{X})^2}, \qquad (3.2) \\ Var(\hat{\beta}_1) &= \frac{\sigma^2}{\sum(X_i - \overline{X})^2} \;, \quad S^2_{\hat{\beta}_1} = \frac{S^2}{\sum(X_i - \overline{X})^2} \;. \end{aligned}$$

In this model setting, to test whether there is any linear relationship between X_t and Y_t, academics test the following hypotheses:

$$H_0^\beta : \beta = \beta_0 \qquad \text{versus} \qquad H_1^\beta : \beta \neq \beta_0 \;, \qquad (3.3)$$

by using the following T^β statistic:

$$T^\beta = \frac{\hat{\beta} - \beta_0}{SE(\hat{\beta})} \;, \qquad (3.4)$$

Readers may refer to Wong and Yue (2024) for more information.

3.2 Model a stationary series with a non-stationary series

To remedy how to test the model of regressing a stationary time series on a non-stationary time series, we first follow Wong and Yue (2024) to define Y_t and X_t as follows:

$$Y_t = \phi Y_{t-1} + \varepsilon_t, \quad \varepsilon_t \overset{iid}{\sim} N(0, \sigma^2_\varepsilon) \;, \quad X_t = X_{t-1} + e_t, \quad e_t \overset{iid}{\sim} N(0, \sigma^2_e) \;, \quad (3.5)$$

with $|\phi| < 1$. We also assume $X_0 = 0, Y_0 = \mu$. Wong and Yue (2024) have shown that regressing a stationary Y_t on a non-stationary X_t may not yield any meaningful outcome and could result in spurious regression if one uses the tests shown in Section 3.1.

4 Remedy

To remedy the issue, first, we propose to use the following remedy:

Remedy 1 *We let*

$$X_t' = (1 - B)X_t = X_t - X_{t-1} \ , \tag{4.1}$$

regress Y_t *on* X_t' *by using the following equation:*

$$Y_t = \alpha + \beta X_t' + u_t \ , \quad t = 1, \cdots, N; \tag{4.2}$$

and make the following conjecture:

Conjecture 4.1 *Regressing* Y_t *on* X_t' *will get meaningful outcome where* Y_t *is defined in Equation* (3.5) *and* X_t' *is defined in Equation* (4.1).

In order to examine the above conjecture, we also examine the following conjecture:

Conjecture 4.2 *Regressing* Y_t *on* X_t' *could not be spurious if one uses the tests from the standard regression model as shown in Section 3.1 where* Y_t *is defined in Equation* (3.5) *and* X_t' *is defined in Equation* (4.1).

However, we believe the following conjecture will hold:

Conjecture 4.3 *The tests from the standard regression model as shown in Section 3.1 is not appropriate to be used for regressing* Y_t *on* X_t' *where* Y_t *is defined in Equation* (3.5) *and* X_t' *is defined in Equation* (4.1).

Thereafter, we propose to use the following remedy:

Remedy 2 *We let*

$$Y_t' = (1 - \hat{\phi}_Y B)Y_t = Y_t - \hat{\phi}_Y Y_{t-1} \ , \tag{4.3}$$

regress Y_t' *on* X_t' *by using the following equation:*

$$Y_t' = \alpha + \beta X_t' + u_t \ , \quad t = 1, \cdots, N; \tag{4.4}$$

and make the following conjecture:

Conjecture 4.4 *Regressing Y_t' on X_t' will get meaningful outcome where both Y_t' and X_t' are defined in Equation* (4.3) *and Equation* (4.1), *respectively.*

In order to examine the above conjecture, we also examine the following conjecture:

Conjecture 4.5 *Regressing Y_t' on X_t' could not be spurious if one uses the tests from the standard regression model as shown in Section 3.1 where Y_t' and X_t' are defined in Equation* (4.3) *and Equation* (4.1), *respectively.*

Last, we propose to use the following remedy:

Remedy 3 *We let*

$$Y_t'' = \frac{Y_t}{1 - \phi_Y B} \quad \textit{and} \quad X_t'' = \frac{X_t}{1 - \phi_X B} \ , \tag{4.5}$$

regress Y_t'' on X_t' by using the following equation:

$$Y_t'' = \alpha + \beta X_t'' + u_t \ , \quad t = 1, \cdots, N; \tag{4.6}$$

and make the following conjecture:

Conjecture 4.6 *Regressing Y_t'' on X_t'' will get meaningful outcome where Y_t'' and X_t'' are defined in Equations* (4.5).

In order to examine the above conjecture, we also examine the following conjecture:

Conjecture 4.7 *Regressing Y_t'' on X_t'' could not be spurious if one uses the tests from the standard regression model as shown in Section 3.1 where Y_t'' and X_t'' are defined in Equations* (4.5).

5 Simulation Setup and Algorithm

To investigate whether all the conjectures stated in this paper hold, we set up a model for simulation and develop algorithms for our simulations in this section, and discuss the simulation results in the next section.

5.1 Model Setup

In this paper, we will modify the setup introduced by Wong and Yue (2024) and briefly state the modified setup here. We note that Wong and Yue (2024) use Equation (3.5) to define both Y_t and X_t and conclude that the statistics cannot be used for the $I0I1$ model. Thus, in this paper, we will not use Equation (3.5) to define both Y_t and X_t in our simulation.

First, we follow Wong and Yue (2024) to use the following four different lengths to control the lengths of time series in our simulation:

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

We then follow Wong and Yue (2024) to use the different values of ϕ in our simulation:

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

5.2 Algorithm

We then modify the simulation procedure introduced by Wong and Yue (2024) to be the algorithm used in this paper. Since X_t and Y_t are generated independently, there is no linear relationship between them. Consequently, under the null hypothesis, H_0^β in (3.3), we expect no systematic rejection, except for random fluctuations at the nominal significance level. For each case defined by the choice of N and ϕ in Section 5.1, we first construct the following algorithm for testing Conjectures 4.1 and 4.2:

Algorithm 1

1. Simulate 10000 independent pairs of time series X_t and Y_t by using the model in Equation (3.5) with coefficients described in Section 5.1.

2. Transform on X_t into its first-difference form according to the remedy approach in Equation (4.1) such that $X'_t = X_t - X_{t-1}$, while keeping Y_t in its original form.

3. Fit the regression model $Y_t = \alpha + \beta X'_t + u_t$ according to Remedy 1, compute $\hat{\beta}$, the test statistic T^{β} in Equation (3.4), and the corresponding p-value.

4. Repeat Steps 2 and 3 for all simulated samples, and record the distribution of $\hat{\beta}$ and p-values.

5. Compute the rejection rate, i.e. the proportion of cases in which $H_0^{\beta} : \beta = 0$ is rejected at the 5% significance level.

We now construct the following algorithm for testing Conjectures 4.4 and 4.5.

Algorithm 2

1. Simulate 10000 independent pairs of time series X_t and Y_t using the model in Equation (3.5) with coefficients described in Section 5.1.

2. Transform both series as follows: $Y'_t = Y_t - \hat{\phi}_Y Y_{t-1}$ and $X'_t = X_t - X_{t-1}$.

3. Fit the regression model $Y'_t = \alpha + \beta X'_t + u_t$ according to Remedy 2, compute $\hat{\beta}$, the test statistic T^{β} in Equation 3.4, and the corresponding p-value.

4. Repeat Steps 2 and 3 for all simulated samples, and record the distribution of $\hat{\beta}$ and p-values.

5. Compute the rejection rate, i.e. the proportion of cases in which $H_0^\beta : \beta = 0$ is rejected at the 5% significance level.

We then construct the following algorithm for testing Conjectures 4.6 and 4.7:

Algorithm 3

1. Simulate 10000 pairs of X_t and Y_t defined in Equation (3.5) with coefficients described in Section 5.1.

2. Let Y_t'' and X_t'' be defined in Equations (4.5).

3. Fit the regression model $Y_t'' = \alpha + \beta X_t'' + u_t$ according to Remedy 3, compute $\hat{\beta}$, the test statistic T^β in 3.4, and the corresponding p-value.

4. Repeat Steps 2 and 3 for all simulated samples, and record the distribution of $\hat{\beta}$ and p-values.

5. Compute the rejection rate, i.e. the proportion of cases in which $H_0^\beta : \beta = 0$ is rejected at the 5% significance level.

The above algorithm helps to examine whether the t statistic in (3.4) for the model in (3.1) follows a Student t-distribution. If X_t and Y_t are unrelated, H_0^β in (3.3) should be rejected 5% of the time if the significance level is 0.05. If $\hat{\beta}$'s follows student t-distribution and the test is perfect, then the rejection rate should be exactly 0.05. If the rejection rate is significantly greater than 0.05, then the results of the t-test cannot be used.

6 Simulation Results

Wong and Yue (2024) use Equation (3.5) to define both Y_t and X_t and found that: (a) when $\phi = 0$, the rejection rate is about 5%; (b) when $\phi > 0$, more than 6% of the independent variables are significant at the 5% level. Moreover, (b1) When N increases, the rejection rate is stable; (b2) When $|\phi|$ increases, the rejection rate increases. In particular, when $\phi = 0.9$, the rejection rate reaches around 60%. Thus, they conclude that the t statistics do not follow the usual distribution under the $I0I1$ model, and standard regression tests cannot be applied directly in this setting.

To remedy the problem, in this paper, we first propose using Remedy 1 by regressing Y_t directly on the transformed series X_t'. The rejection rates for this setting are reported in Table 6.1.

Table 6.1: Rejection Rate in the simulation for various ϕ and N when regressing Y_t on X_t'.

Case	Coefficients	N=100	N=500	N=1000	Average
(a)	$\phi = 0$	0.0496	0.0487	0.0523	0.0502
(b)	$\phi = 0.1$	0.0385	0.0384	0.0406	0.0392
(c)	$\phi = 0.3$	0.0196	0.0190	0.0198	0.0195
(d)	$\phi = 0.5$	0.0074	0.0071	0.0057	0.0067
(e)	$\phi = 0.7$	0.0011	0.0001	0.0004	0.0005
(f)	$\phi = 0.9$	0.0000	0.0000	0.0000	0.0000

Note: $X_t' = (1 - B)X_t = X_t - X_{t-1}$. Readers may read Remedy 1 for more information.

Table 6.1 paints a strikingly different picture. When $\phi = 0$, Y_t consists purely of white noise, and the rejection rate is close to the nominal 5% level. However, as ϕ increases, the rejection rate declines monotonically. By the time $\phi \geq 0.7$, the rejection rate falls to nearly zero, and for $\phi = 0.9$, the test never rejects the null at all across 10,000 replications. This pattern suggests that regressing Y_t on

X_t' systematically over-rejects as persistence in ϕ increases. The test no longer has the correct size and becomes increasingly conservative. In other words, although X_t' is stationary, using it to explain the original Y_t does not restore valid inference. These results support Conjectures 4.1 and 4.2 to the extent that the regression does not generate spurious significance; however, they also confirm Conjecture 4.3, which posits that the standard t test is not appropriate in this case. In fact, the dramatic under-rejection indicates that the classical inference framework breaks down, and new testing approaches may be needed for the Y_t and X_t' specification. The results in Table 6.1 suggest that though using Remedy 1 to regress Y_t on $X_t' = X_t - X_{t-1}$ is a good remedy to correct the spurious problem, it results in a smaller size. Thus, this concludes that Remedy 1 is not a good remedy. We will explore the issue in the future.

We then propose using Remedy 2 by regressing $Y_t' = (1 - \hat{\phi}_Y B)Y_t = Y_t - \hat{\phi}_Y Y_{t-1}$ on $X_t' = (1-B)X_t = X_t - X_{t-1}$ and examine whether this transformation restores the validity of the test. The rejection rates under this setting are reported in Table 6.2:

Table 6.2: Rejection Rate in the simulation for various ϕ and N when regressing Y_t' on X_t'.

Case	Coefficients	N=100	N=500	N=1000	Average
(a)	$\phi = 0$	0.0488	0.0486	0.0485	0.0486
(b)	$\phi = 0.1$	0.0490	0.0473	0.0478	0.0480
(c)	$\phi = 0.3$	0.0482	0.0510	0.0473	0.0488
(d)	$\phi = 0.5$	0.0483	0.0480	0.0513	0.0492
(e)	$\phi = 0.7$	0.0480	0.0458	0.0485	0.0474
(f)	$\phi = 0.9$	0.0446	0.0473	0.0535	0.0485
Overall Average					**0.0484**

Note: $Y_t' = (1 - \hat{\phi}_Y B)Y_t = Y_t - \hat{\phi}_Y Y_{t-1}$ and $X_t' = (1 - B)X_t = X_t - X_{t-1}$. Readers may read Remedy 2 for more information.

Table 6.2 shows that the rejection rates remain extremely close to the nominal 5% level across all cases, regardless of the value of ϕ or the sample size N. Even at the extreme case $\phi = 0.9$, the rejection rate is 4.9%, which is nearly identical to the theoretical benchmark. These results provide strong evidence that the remedy transformation successfully eliminates the spurious regression problem. The t test based on the standard regression model behaves as expected, and the distribution of the test statistic appears valid. This supports both Conjecture 4.4 (that regressing Y_t' on X_t' yields meaningful outcomes) and Conjecture 4.5 (that the regression is not spurious under the standard test framework). The results in Table 6.2 suggest that using Remedy 2 to regress $Y_t' = (1 - \hat{\phi}_Y B)Y_t = Y_t - \hat{\phi}_Y Y_{t-1}$ on $X_t' = (1 - B)X_t = X_t - X_{t-1}$ is not only a good remedy to correct the spurious problem, but also a good remedy to keep the size to be close to the theoretical benchmark. Thus, this concludes that Remedy 2 is a very good remedy.

Table 6.3: Rejection Rate in the simulation for various ϕ and N when regressing Y_t'' on X_t'.

ϕ_Y	N=100	N=500	N=1000	Average
$\phi_Y = 0$	0.0498	0.0472	0.0514	0.0495
$\phi_Y = 0.1$	0.0506	0.0483	0.0516	0.0502
$\phi_Y = 0.3$	0.0489	0.0495	0.0515	0.05
$\phi_Y = 0.5$	0.05	0.0472	0.0518	0.0497
$\phi_Y = 0.7$	0.0456	0.0501	0.0509	0.0489
$\phi_Y = 0.9$	0.0481	0.0506	0.0495	0.0494
Overall Average				**0.0496**

Note: $Y_t'' = \frac{Y_t}{1-\phi_Y B}$ and $X_t' = \frac{X_t}{1-\phi_X B}$. Readers may read Remedy 3 for more information.

Similar to the results in Table 6.2, Table 6.3 also shows that the rejection rates remain extremely close to the nominal 5% level across all cases, regardless of the value of ϕ or the sample size N. Even at the extreme case $\phi = 0.9$, the rejection rate is 4.9%, which is nearly identical to the theoretical benchmark. These results provide strong evidence that the remedy transformation successfully eliminates the spurious regression problem. The t test based on the standard regression model behaves as expected, and the distribution of the test statistic appears valid. This supports both Conjecture 4.6 (that regressing Y_t'' on X_t'' yields meaningful outcomes) and Conjecture 4.7 (that the regression is not spurious under the standard test framework). The results in Table 6.3 suggest that using Remedy 3 to regress $Y_t'' = \frac{Y_t}{1-\phi_Y B}$ on $X_t'' = \frac{X_t}{1-\phi_X B}$ is not only a good remedy to correct the spurious problem, but also a good remedy to keep the size to be close to the theoretical benchmark. Thus, this concludes that Remedy 3 is a very good remedy.

Now, we examine between Remedy 2 and Remedy 3, which one is better. Since the overall average of the last column in Table 6.2 is 0.0484, while the overall average of the last column in Table 6.3 is 0.0496, which is closer to the theoretical benchmark. Thus, we conclude that Remedy 3 is slightly better according to our simulation.

7 Conclusion and Future Study

Based on the foundational work of Granger and Newbold (1974) to address the issue of regressing a stationary time series, Y_t, on a non-stationary time series, X_t, (we call it the I0I1 model), many papers in the literature report results of the I0I1 model. For example, Singh et al. (2011) regressed stock returns on GDP. However, very few studies have examined whether there are any problems with using the model. Recently, Wong and Yue (2024) conducted a simulation and found that regressing a stationary time series, Y_t, on a non-stationary time series, X_t, could be spurious. They then show 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 the I0I0 model) do not have any asymptotic distribution with both mean and variance tending to infinity and conclude that the statistics cannot be used for the $I0I1$ model. To bridge the gap in the literature, in this paper, we introduced three remedies to overcome the limitations of the statistics.

Our first proposed remedy (Remedy 1) is to recommend academics and practitioners regress Y_t on $X_t' = X_t - X_{t-1}$ (see Equation (4.1)). Our simulation results (as shown in Table 6.1) suggest that using Remedy 1 could correct the spurious problem, but it results in getting a much smaller size. Thus, this concludes that Remedy 1 is not a good remedy.

Our second proposed remedy (Remedy 2) is to recommend academics and practitioners regress $Y_t' = (1 - \hat{\phi}_Y B)Y_t$ (see Equation (4.3)) on $X_t' = X_t - X_{t-1}$ (see Equation (4.1)). Our simulation results (as shown in Table 6.2) suggest that Remedy 2 is not only a good remedy to correct the spurious problem, but also a good remedy to keep the size close to the theoretical benchmark. Thus, we conclude that Remedy 2 is a very good remedy.

Our third proposed remedy (Remedy 3) is to recommend academics and practitioners regress $Y_t'' = \frac{Y_t}{1-\phi_Y B}$ on $X_t'' = \frac{X_t}{1-\phi_X B}$ (see Equation (4.5)). Similarly, our simulation results (as shown in Table 6.3) suggest that Remedy 3 is also a good remedy to correct the spurious problem as well as keep the size close to the theoretical benchmark. Thus, we conclude that Remedy 3 is also a very good remedy. To examine between Remedy 2 and Remedy 3, which one is better, we conducted simulation and found that the overall average of the rejection rate (see Table 6.2) by using Remedy 2 is 0.0484, while the overall average of the rejection rate (see Table 6.3) by using Remedy 3 is 0.0496, which is closer to the theoretical benchmark. Thus, we conclude that Remedy 3 is slightly better according to our simulation.

There could be some limitations of our paper. For example, the models used in our paper is very simple. Thus, extensions of our paper include studying more complicated models. Second, our paper only provided the results of the simulation but did not develop any theory for the issue. Thus, extensions of our paper include developing some theories on the issue, see, for example, Wong and Pham (2022) and Wong and Pham (2024*b*). We also note that Wong and Pham (2022) and Wong and Pham (2023) have shown that the test from the standard regression model could make significant regression with autoregressive noise become insignificant for small sample. Academics could incorporate the idea of this paper to extend the work from Wong and Pham (2022) and Wong and Pham (2023).

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.

Cheng, Y., Hui, Y., Liu, S. and Wong, W.-K. (2022), 'Could significant regression be treated as insignificant: An anomaly in statistics?', *Communications in Statistics: Case Studies, Data Analysis and Applications* **8**(1), 133–151.

Cheng, Y., Hui, Y., McAleer, M. and Wong, W.-K. (2021), 'Spurious relationships for nearly non-stationary series', *Journal of Risk and Financial Management* **14**(8), 366.

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.

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.

Wong, W.-K., Cheng, Y. and Yue, M. (2024), 'Could regression of stationary series be spurious?', *Asia-Pacific Journal of Operational Research* .

Wong, W.-K. and Pham, M. T. (2022), 'Could the test from the standard regression model could make significant regression with autoregressive noise become insignificant?', *The International Journal of Finance* **34**, 1–18.

Wong, W.-K. and Pham, M. T. (2023), 'Could the test from the standard regression model could make significant regression with autoregressive $\mathrm{y_t}$ and $\mathrm{x_t}$ become insignificant?', *The International Journal of Finance* **35**, 1–19.

Wong, W.-K. and Pham, M. T. (2024*a*), 'Could the correlation of a stationary series with a non-stationary series obtain meaningful outcomes?', *Annals of Financial Economics* . Forthcoming.

Wong, W.-K. and Pham, M. T. (2024*b*), 'How to model a simple stationary series with a non-stationary series?', *The International Journal of Finance* . forthcoming.

Wong, W.-K. and Yue, M. (2024), 'Could regressing a stationary series on a non-stationary series obtain meaningful outcomes?', *Annals of Financial Economics* **19**(03), 2450011.