Kagawa University Economic Review
Vol. 80, No. 2, September 2007, 37－52

Nonsense Regressions in Econometrics, I(1) with drift vs. Trend Stationary.

Hiroyuki Hisamatsu

Abstract

In this paper we consider the spurious or nonsense regression phenomenon where the DGPs of the regressor and the regressand are I(1) with drift vs. trend stationary and I(0) vs. I(1) with drift, and all of these patterns have first order autoregressive errors. We derive the asymptotic distributions or probability limits of the OLS estimator, the conventional significance t test, R^2 and DW statistics. In these cases it is found that the spurious or nonsense regression phenomenon occurs and we examine the effect of drifts and AR(1) coefficients of the errors of regressor or regressand to the asymptotic distributions of the OLS estimator and the associated test statistics.

Keywords: nonsense regression, I(1) with drift, trend stationary

JEL Classification: C12

1 Introduction

Spurious regression or nonsense regression were studied by Granger and Newbold(1974) using Monte Carlo simulation and they have shown that the spurious regression phenomenon occurs when independent random walks are regressed on one another. This phenomenon was also analyzed analytically by Phillips(1986). Durlauf and Phillips(1988) studied the problem of spurious detrending.

Entorf(1997) studied the phenomenon of the two independent random walks with non-zero drifts and he has shown that in the I(1) with drift model it causes the convergence of the OLS to the ratio of drifts. Comparing the results of Phillips(1986)'s without drift case, non-zero drift asymptotics differ in several ways. Tsay and Chung(2000) extended the theoretical analysis of the spurious regression and spurious detrending from the usual I(1) processes to the long memory fractionally integrated processes. They found that the spurious effects occur. He and Maekawa(2001) studied the similar phenomenon of Granger causality test for two independent variables. Kim, Lee and Newbold(2004) considered the situation where the deterministic components of the processes generating individual series are linear trends and they are independent I(0) processes. It is shown that when in these cases nonsense regression phenomenon occurs. Noriega and Santaularia(2005) considered the phenomenon between broken trend variables.

This paper is organized as follows: Section 2 presents the 4 different patterns of the DGPs of regressor and regressand and derives the asymptotic distributions or probability limits of the OLS estimator, the conventional significance t test, R^2 and DW statistics of the regression analysis. Section 3 examines the small sample properties of these statistics by Monte Carlo experiments. Section 4 provides some concluding remarks.

2 Models and the asymptotics

We assume all statistical inferences are based on the hypothesis that $y_t = \alpha + \gamma x_t + u_t, t = 1, 2, ..., T$ is a correct statistical model. A researcher might run an OLS regression of y_t on a constant and x_t. The estimated regression equation is given by $y_t = \hat{\alpha} + \hat{\gamma} x_t + e_t$ where e_t is a residual. We derive the asymptotic distributions or probability limits of the OLS estimator and the associated test statistics of the regression analysis for 4 different patterns of the DGPs of regressor and regressand in line with Kim, Lee and Newbold(2004) and Entorf(1997). We consider the following data generating processes(DGP) of $\{y_t\}$ and $\{x_t\}$.

DGP1: I(1) with drift vs. trend stationary

$$y_t = \mu_y + \beta_y y_{t-1} + u_{yt}, \quad u_{yt} = \phi_y u_{yt-1} + \varepsilon_{yt}, \quad |\phi_y| < 1, \tag{1}$$

$$x_t = \mu_x + \beta_x t + u_{xt}, \quad u_{xt} = \phi_x u_{xt-1} + \varepsilon_{xt}, \quad |\phi_x| < 1, \tag{2}$$

where$\beta_y = 1$, $\beta_x \neq 1$, $\varepsilon_{yt} \sim iid(0, \sigma_y^2)$, $\varepsilon_{xt} \sim iid(0, \sigma_x^2)$, ε_{yt} and ε_{xt} are independent.

Theorem 1 *If the DGPs of $\{y_t\}$ and $\{x_t\}$ are given by the equations (1) and(2), the least squares estimator $\widehat{\gamma}$ converges in*

$$\widehat{\gamma} = \frac{\sum_{t=1}^{T}(y_t - \overline{y})(x_t - \overline{x})}{\sum_{t=1}^{T}(x_t - \overline{x})^2} \xrightarrow{p} \mu_y/\beta_x,$$

and the asymptotic distribution of the t-statistics $t_{\widehat{\gamma}}$ is given by the functionals of the Brownian motion as

$$\begin{aligned}(1/T)\,t_{\widehat{\gamma}} &= (1/T)\widehat{\gamma}/\left\{\left(\sum_{t=1}^{T} e_t^2/(T-2)\right)/\sum_{t=1}^{T}(x_t - \overline{x})^2\right\}^{1/2} \\ &\xrightarrow{p} \mu_y/\left\{12\sigma_y^2/(1-\phi_y^2)\int_0^1 W_y(r)^2 dr\right\}^{1/2}.\end{aligned}$$

The asymptotic distribution of the DW is given as

$$TDW = T\left\{\sum_{t=2}^{T}(e_t - e_{t-1})^2/\sum_{t=1}^{T} e_t^2\right\} \xrightarrow{p}$$

$$\left[\sigma_y^2/(1-\phi_y^2) + (\mu_y/\beta_x)^2\left\{2\sigma_x^2/(1+\phi_x)\right\}\right]\left\{\sigma_y^2/(1-\phi_y^2)\int_0^1 W_y(r)^2 dr\right\}^{-1},$$

and those of the R^2 (or equivalently to those of the $\overline{R}^2$) as

$$\begin{aligned}T\left(1-R^2\right) &= \left(\sum_{t=1}^{T} e_t^2/T^2\right)\left\{\left(1/T^3\right)\sum_{t=1}^{T}(y_t - \overline{y})^2\right\}^{-1} \\ &\xrightarrow{p} \left(12/\mu_y^2\right)\left\{\sigma_y^2/(1-\phi_y^2)\right\}\int_0^1 W_y(r)^2 dr. \blacksquare\end{aligned}$$

Proof of Theorem 1: We have $y_t = \mu_y + \beta_y y_{t-1} + u_{yt} = \mu_y t + y_0 + \sum_{j=1}^{t} u_{yj} = \mu_y t + y_0 + S_{yt}$.We recognize that, if $\mu_y = \beta_x$, $\{y_t\}$ and $\{x_t\}$ have a common trend. The results $(1/T^2)\sum_{t=1}^{T} y_t \xrightarrow{p} \mu_y/2$, $(1/T^2)\sum_{t=1}^{T} x_t \xrightarrow{p} \beta_x/2$, $(1/T^3)\sum_{t=1}^{T} y_t x_t \xrightarrow{p} \mu_y\beta_x/3$ lead directly to$(1/T^3)\sum_{t=1}^{T}(x_t-\overline{x})^2 \xrightarrow{p} \beta_x^2/12$, $(1/T^3)\sum_{t=1}^{T}(y_t-\overline{y})(x_t-\overline{x}) \xrightarrow{p} \mu_y\beta_x/12$. We can rewrite the DGP of $\{y_t\}$ as follows.$y_t = \mu_y t + y_0 + S_{yt} = (y_0 - \gamma\mu_x) + \gamma x_t + (S_{yt} - \gamma u_{xt}) = \alpha + \gamma x_t + u_t$ where $\gamma = \mu_y/\beta_x$. The residual can be written as $e_t = S_{yt} - \widehat{\gamma} u_{xt}$. We have

$$(1/T^2)\sum_{t=1}^{T} e_t^2 = (1/T^2)\sum_{t=1}^{T} (S_{yt} - \widehat{\gamma} u_{xt})^2 \xrightarrow{p} \sigma_{uy}^2 \int_0^1 W_y(r)^2 dr$$

where $\sigma_{uy}^2 = \sigma_y^2/(1-\phi_y^2)$ and $W_y(r)$ is a Wiener process. Using these results we can derive these asymptotic distributions.

Remark 1: In this case the OLS estimator $\widehat{\gamma}$ converges in probability to the ratio of drifts. The t-statistics $t_{\widehat{\gamma}}$ diverges with T and thus we have a nonstandard asymptotic distribution of $t_{\widehat{\gamma}}$ scaled by T^{-1} as stated above. The *DW* statistics diverges with T and thus approaches 0. The R^2 (or $\overline{R}^2$) approaches unity with speed T.

In a similar way, we can derive those for the case that the following DGPs of $\{y_t\}$ and $\{x_t\}$.

DGP 2: Trend stationary vs. I(1) with drift

$$y_t = \mu_y + \beta_y t + u_{yt}, \quad u_{yt} = \phi_y u_{yt-1} + \varepsilon_{yt}, \quad |\phi_y| < 1, \tag{3}$$
$$x_t = \mu_x + \beta_x x_{t-1} + u_{xt}, \quad u_{xt} = \phi_x u_{xt-1} + \varepsilon_{xt}, \quad |\phi_x| < 1, \tag{4}$$

where$\beta_y \neq 1$, $\beta_x = 1$, $\varepsilon_{yt} \sim iid(0,\sigma_y^2)$, $\varepsilon_{xt} \sim iid(0,\sigma_x^2)$, ε_{yt} and ε_{xt} are independent.

Theorem 2 *If the DGPs of $\{y_t\}$ and $\{x_t\}$ are given by the equations (3) and (4), the least squares estimator $\widehat{\gamma}$ converges in*

$$\widehat{\gamma} = \frac{\sum_{t=1}^{T}(y_t - \overline{y})(x_t - \overline{x})}{\sum_{t=1}^{T}(x_t - \overline{x})^2} \xrightarrow{p} \beta_y/\mu_x,$$

and the asymptotic distribution of the t-statistics $t_{\widehat{\gamma}}$ is given by the functionals of the Brownian motion as

$$(1/T)\, t_{\widehat{\gamma}} = (1/T)\,\widehat{\gamma}/\left[\left\{\sum_{t=1}^{T} e_t^2/(T-2)\right\}/\sum_{t=1}^{T}(x_t - \overline{x})^2\right]^{1/2}$$
$$\xrightarrow{p} \mu_x/\left\{12\sigma_x^2/(1-\phi_x^2)\int_0^1 W_x(r)^2 dr\right\}^{1/2}.$$

The asymptotic distribution of the DW is given as

$$T \cdot DW = T\left\{\sum_{t=2}^{T}(e_t - e_{t-1})^2/\sum_{t=1}^{T} e_t^2\right\} \xrightarrow{p}$$
$$\left[2\sigma_y^2/(1+\phi_y) + (\beta_y/\mu_x)^2\left\{\sigma_x^2/\left(1-\phi_x^2\right)\right\}\right]$$
$$\times\left[(\beta_y/\mu_x)^2\left\{\sigma_x^2/(1-\phi_x^2)\right\}\int_0^1 W_x(r)^2 dr\right]^{-1},$$

and those of the R^2 (or $\overline{R}^2$) as

$$T\left(1-R^2\right) = \left(\sum_{t=1}^{T} e_t^2/T^2\right)\left\{\left(1/T^3\right)\sum_{t=1}^{T}(y_t - \overline{y})^2\right\}^{-1}$$
$$\xrightarrow{p} \left(12/\mu_x^2\right)\left\{\sigma_x^2/(1-\phi_x^2)\right\}\int_0^1 W_x(r)^2 dr. \blacksquare$$

Proof of Theorem 2:

We have $x_t = \mu_x t + x_0 + S_{xt}$ where $S_{xt} = \sum_{j=1}^{t} u_{xj}$. The results

$$(1/T^2)\sum_{t=1}^{T} x_t \xrightarrow{p} \mu_x/2,\ (1/T^2)\sum_{t=1}^{T} y_t \xrightarrow{p} \beta_y/2,$$
$$(1/T^3)\sum_{t=1}^{T} y_t x_t \xrightarrow{p} \mu_x\beta_y/3$$

lead directly to

$$(1/T^3)\sum_{t=1}^{T}(x_t-\overline{x})^2 \xrightarrow{p} \mu_x^2/12,\ (1/T^3)\sum_{t=1}^{T}(y_t-\overline{y})^2 \xrightarrow{p} \beta_y^2/12,$$

$$(1/T^3)\sum_{t=1}^{T}(y_t-\overline{y})(x_t-\overline{x}) \xrightarrow{p} \mu_x\beta_y/12.$$

We can rewrite $y_t = \mu_y + \beta_y t + u_{yt} = \mu_y + \beta_y\left\{(x_t - x_0 - S_{xt})/\mu_x\right\} + u_{yt} = (\mu_y - \gamma x_0) + \gamma x_t + (u_{yt} - \gamma S_{xt}) = \alpha + \gamma x_t + u_t$ where $\gamma = \beta_y/\mu_x$. The residual can be written as $e_t = u_{yt} - \widehat{\gamma} S_{xt}$. We have $(1/T^2)\sum_{t=1}^{T} e_t^2 = (1/T^2)\sum_{t=1}^{T}(u_{yt} - \widehat{\gamma} S_{xt})^2 = (1/T^2)\sum_{t=1}^{T}\{u_{yt}^2 - 2\widehat{\gamma} u_{yt} S_{xt} + \widehat{\gamma}^2 S_{xt}^2\} \longrightarrow \gamma^2\sigma_{ux}^2 \int_0^1 W_x(r)^2 dr$,where $\sigma_{ux}^2 = \sigma_x^2/(1-\phi_x^2)$ and $W_x(r)$ is a Wiener process. Using these results we can derive these asymptotic distributions.

Remark 2: Our Theorem 1 and 2 are similar with the Theorem 1 of Kim, Lee and Newbold(2004) in the sense that $\widehat{\gamma}$ converges in probability to the ratio of coefficients of the trend. But the asymptotics of $t_{\widehat{\gamma}}$ is slightly different. The former has a limiting distribution but the latter has a probability limit.

Remark3: The asymptotic results of the DW and R^2(or $\overline{R}^2$) of the Kim,et al. can be derived as follows.

$$DW \xrightarrow{p} \left\{2\sigma_y^2/\left(1+\phi_y\right) + \gamma^2 \cdot 2\sigma_x^2/(1+\phi_x)\right\}$$
$$/\left\{\sigma_y^2/(1-\phi_y^2) + \gamma^2 \cdot \sigma_x^2/(1-\phi_x^2)\right\},$$

and

$$T^2\left(1-R^2\right) \xrightarrow{p} \left\{\sigma_y^2/(1-\phi_y^2) + \gamma^2 \cdot \sigma_x^2/(1-\phi_x^2)\right\} / \left(\beta_y^2/12\right).$$

The DW does not converge in probability to 0 and the R^2 converges to 1 faster than our cases.

We consider the following DGPs of $\{y_t\}$ and $\{x_t\}$.

DGP3: I(0) vs. I(1) with drift

$$y_t = \mu_y + \beta_y t + u_{yt}, \quad u_{yt} = \phi_y u_{yt-1} + \varepsilon_{yt}, \quad |\phi_y| < 1, \tag{5}$$
$$x_t = \mu_x + \beta_x x_{t-1} + u_{xt}, \quad u_{xt} = \phi_x u_{xt-1} + \varepsilon_{xt}, \quad |\phi_x| < 1, \tag{6}$$

where $\beta_y = 0, \beta_x = 1$, $\varepsilon_{yt} \sim iid(0, \sigma_y^2)$, $\varepsilon_{xt} \sim iid(0, \sigma_x^2)$, and ε_{yt} and ε_{xt} are independent.

Theorem 3 *If the DGPs of $\{y_t\}$ and $\{x_t\}$ are given by the equations (5) and (6), the asymptotic distribution of the least squares estimator $\widehat{\gamma}$ is given as*

$$\begin{aligned} T^{3/2}\widehat{\gamma} &= \left\{T^{-3/2}\sum_{t=1}^{T}(y_t-\overline{y})(x_t-\overline{x})\right\} / \left\{T^{-3}\sum_{t=1}^{T}(x_t-\overline{x})^2\right\} \\ &\xrightarrow{d} N\left(0,\ 12\sigma_y^2/\left\{\mu_x^2\left(1-\phi_y\right)^2\right\}\right), \end{aligned}$$

and the asymptotic distribution of the t-statistics $t_{\widehat{\gamma}}$ is given as

$$\begin{aligned} t_{\widehat{\gamma}} &= \widehat{\gamma}/\left\{\left(\sum_{t=1}^{T} e_t^2/(T-2)\right)/\sum_{t=1}^{T}(x_t-\overline{x})^2\right\}^{1/2} \\ &\xrightarrow{d} N\left(0,\ \left(1+\phi_y\right)/\left(1-\phi_y\right)\right). \end{aligned}$$

The asymptotic distribution of the DW is given as

$$DW = \left\{ \sum_{t=2}^{T} (e_t - e_{t-1})^2 / T \right\} \left(\sum_{t=1}^{T} e_t^2 / T \right)^{-1}$$

$$\xrightarrow{p} \left\{ 2\sigma_y^2/(1+\phi_y) \right\} \left\{ \sigma_y^2/(1-\phi_y^2) \right\}^{-1} = 2\left(1-\phi_y\right),$$

and those of the R^2 *(or* $\overline{R}^2$*) as*

$$\begin{aligned} R^2 &= 1 - \left(\sum_{t=1}^{T} e_t^2/T \right) \left\{ (1/T) \sum_{t=1}^{T} (y_t - \overline{y})^2 \right\}^{-1} \\ &\xrightarrow{p} 1 - \left\{ \sigma_y^2/(1-\phi_y^2) \right\} \left\{ \sigma_y^2/(1-\phi_y^2) \right\}^{-1} = 0. \blacksquare \end{aligned}$$

Proof of Theorem 3: We can rewrite as $x_t = \mu_x + \beta_x x_{t-1} + u_{xt} = \mu_x t + x_0 + \sum_{j=1}^{t} u_{xj} = \mu_x t + x_0 + S_{xt}$. It is straightforward that

$$\left(T^{-3/2}\right) \sum_{t=1}^{T} (y_t - \overline{y})(x_t - \overline{x}) = \mu_x A' \begin{pmatrix} T^{-1/2} \sum_{t=1}^{T} u_{yt} \\ T^{-3/2} \sum_{t=1}^{T} u_{yt} t \end{pmatrix} + o_p(1)$$
$$\xrightarrow{d} N\left(0,\ \mu_x^2 \sigma_y^2 / \left\{ 12\left(1-\phi_y\right)^2 \right\}\right), \text{ where } A = (-1/2, 1)'.$$

On the other hand, $T^{-3} \sum_{t=1}^{T} (x_t - \overline{x})^2 \xrightarrow{p} \mu_x^2/12$. Using these results we can derive these asymptotic distributions.

Remark 4: In this case the normalized OLS estimator $\widehat{\gamma}$ with $T^{3/2}$ have an asymptotic normal distribution. The t-statistics $t_{\widehat{\gamma}}$ have also an asymptotic normal distribution. The DW statistics converges in probability to $2\left(1-\phi_y\right)$. The R^2 approaches 0.

We consider the following DGPs of $\{y_t\}$ and $\{x_t\}$.

DGP 4: I(1) with drift vs. I(0)

$$y_t = \mu_y + \beta_y y_{t-1} + u_{yt}, \quad u_{yt} = \phi_y u_{yt-1} + \varepsilon_{yt}, \quad |\phi_y| < 1, \tag{7}$$
$$x_t = \mu_x + \beta_x t + u_{xt}, \quad u_{xt} = \phi_x u_{xt-1} + \varepsilon_{xt}, \quad |\phi_x| < 1, \tag{8}$$

where $\beta_y = 1$ and $\beta_x = 0, \varepsilon_{yt} \sim iid(0, \sigma_y^2)$, $\varepsilon_{xt} \sim iid(0, \sigma_x^2)$, and ε_{yt} and ε_{xt} are independent.

Theorem 4 *If the DGPs of $\{y_t\}$ and $\{x_t\}$ are given by the equations (7) and (8), the asymptotic distribution of the least squares estimator $\widehat{\gamma}$ is given as*

$$T^{-1/2}\widehat{\gamma} \xrightarrow{d} N\left(0, \ \mu_y^2 / \left(12\sigma_x^2\right) \cdot (1+\phi_x)^2\right),$$

and the asymptotic distribution of the t-statistics $t_{\widehat{\gamma}}$ is given as

$$t_{\widehat{\gamma}} \xrightarrow{d} N\left(0, \ (1+\phi_x) / (1-\phi_x)\right).$$

The asymptotic distribution of the DW is given as

$$\begin{aligned} TDW &= \left\{\sum_{t=2}^{T} (e_t - e_{t-1})^2 / T^2\right\} \left(\sum_{t=1}^{T} e_t^2 / T^3\right)^{-1} \\ &\xrightarrow{p} \left[24\sigma_x^2 / \left\{\mu_y^2 (1+\phi_x)\right\}\right] \cdot \xi^2 \end{aligned}$$

where $\xi \sim N\left(0, \mu_y^2 / \left(12\sigma_x^2\right) \cdot (1+\phi_x)^2\right)$, and those of the R^2 (or $\overline{R}^2$) as

$$\begin{aligned} R^2 &= 1 - \left(\sum_{t=1}^{T} e_t^2 / T^3\right) \left\{\left(1/T^3\right) \sum_{t=1}^{T} (y_t - \overline{y})^2\right\}^{-1} \\ &\xrightarrow{p} 1 - \left\{\mu_y^2 / 12\right\} \left\{\mu_y^2 / 12\right\}^{-1} = 0. \blacksquare \end{aligned}$$

Proof of Theorem 4: We can rewrite the DGP of $\{y_t\}$ as $y_t = \mu_y + \beta_y y_{t-1} +$

$u_{yt} = \mu_y t + y_0 + \sum_{j=1}^{t} u_{yj} = \mu_y t + y_0 + S_{yt}$. We have

$$\left(T^{-3/2}\right)\sum_{t=1}^{T}(y_t-\overline{y})(x_t-\overline{x}) = \mu_y A' \begin{pmatrix} T^{-1/2}\sum_{t=1}^{T} u_{xt} \\ T^{-3/2}\sum_{t=1}^{T} u_{xt}t \end{pmatrix} + o_p(1)$$

$$\xrightarrow{d} N\left(0,\ \mu_y^2\sigma_x^2/\left\{12\left(1-\phi_x\right)^2\right\}\right),\ \text{where } A = (-1/2, 1)'.$$

On the other hand,

$$\begin{aligned} T^{-1}\sum_{t=1}^{T}\left(x_t-\overline{x}\right)^2 &= T^{-1}\sum_{t=1}^{T}\left(u_{xt}-\sum u_{xt}/T\right)^2 \\ \xrightarrow{p} \sigma_{ux}^2 &= \sigma_x^2/\left(1-\phi_x^2\right). \end{aligned}$$

and

$$\left\{\sum_{t=1}^{T} e_t^2/\left(T-2\right)\right\}/T^2 = \sum_{t=1}^{T}(y_t-\overline{y})^2/T^3 + o_p(1) \xrightarrow{p} \mu_y^2/12,$$

$$\begin{aligned} \left(T^{-2}\right)\sum_{t=2}^{T}\left(e_t-e_{t-1}\right)^2 &= \left(T^{-2}\right)\hat{\gamma}^2\sum_{t=2}^{T}\left(u_{xt}-u_{xt-1}\right)^2 + o_p(1) \\ &\xrightarrow{p} \left\{2\sigma_x^2/\left(1+\phi_x\right)\right\}\cdot\xi^2, \end{aligned}$$

where $\xi \sim N\left(0, \mu_y^2/\left(12\sigma_x^2\right)\cdot\left(1+\phi_x\right)^2\right)$. Using these results we can derive these asymptotic distributions.

Remark 5: Theorem 3 and 4 imply that as the AR parameter ϕ approaches unity, the asymptotic rejection rate of the t-statistics $t_{\hat{\gamma}}$ becomes larger. Therefore, the phenomenon of the nonsense regression depends on closeness of ϕ to unity. Our Theorem 3 is similar with the upper case of Theorem2 of Kim, et al in the sense that $T^{3/2}\widehat{\gamma} \xrightarrow{d} Normal$ and $t_{\widehat{\gamma}} \xrightarrow{d} Normal.$

Remark 6: To compare the asymptotic results of the DW and R^2(or $\overline{R}^2$) , we derive those for Kim, et al. $DW = \left\{\sum_{t=2}^{T}(e_t - e_{t-1})^2 / \sum_{t=1}^{T} e_t^2\right\} \xrightarrow{p} 2(1-\phi_y), 1-R^2 = \left(\sum_{t=1}^{T} e_t^2/T\right)\left\{\sum_{t=1}^{T}(y_t-\overline{y})^2/T\right\}^{-1} \xrightarrow{p} \{\sigma_y^2/(1-\phi_y^2)\}\{\sigma_y^2/(1-\phi_y^2)^{-1}\} = 1$. These results are similar with our Theorem 3.

Remark 7: On the other hand, our Theorem 4 is similar with the lower case of Theorem 2 of Kim, et al. in the sense that $T^{-1/2}\widehat{\gamma} \xrightarrow{d} Normal$ and $t_{\widehat{\gamma}} \xrightarrow{d} Normal$. The asymptotic results of the DW and R^2(or $\overline{R}^2$) *of* Kim et al. can be derived as follows.

$$\begin{aligned} TDW &= \left\{\sum_{t=2}^{T}(e_t - e_{t-1})^2 / T^2\right\}\left(\sum_{t=1}^{T} e_t^2/T^3\right)^{-1} \\ &\xrightarrow{p} \left[24\sigma_x^2 / \left\{\beta_y^2(1+\phi_x)\right\}\right]\cdot\xi^2 \end{aligned}$$

where$\xi \sim N\left(0, \beta_y^2/\left(12\sigma_x^2\right)\cdot(1+\phi_x)^2\right)$,

$$\begin{aligned} 1-R^2 &= \left(\sum_{t=1}^{T} e_t^2/T^3\right)\left\{\sum_{t=1}^{T}(y_t-\overline{y})^2/T^3\right\}^{-1} \\ \xrightarrow{p} (\beta_y^2/12)/(\beta_y^2/12) &= 1. \end{aligned}$$

These results are similar with our Theorem 4.

3 Simulated Results

In order to analyze the finite sample properties of the t-statistics, DW ratio and $\overline{R}^2$, we perform some Monte Carlo experiments for T=50, T=1000 and 1000 replications. The two error terms $\{\varepsilon_{yt}\}$ and $\{\varepsilon_{xt}\}$ are drawn from N(0,1) and we set $y_0 = 0$, $x_0 = 0$. The indices of the following tables are as follows. From the left, the number of rejection ($|t_{\widehat{\gamma}}| \geq 1.96$), averages of DW, averages of $\overline{R}^2$, the number of times of $\overline{R}^2 > 0.7$, the number of times of $DW < 1.22$, respectively. The upper row is for T=50 and the lower row is for T=1000.

Table 3.1: Results for DGPs (1), (2) where $\beta_y = 1$.

β_x	ϕ_y	ϕ_x	$\lvert t_{\hat{\gamma}}\rvert \geq 1.96$	DW	$\overline{R}^2$	$\overline{R}^2 > 0.7$	$DW < 1.22$
0.2	0.0	0.9	997	0.377	0.684	560	999
			1000	0.091	0.996	1000	1000
0.2	0.9	0.0	915	0.619	0.572	467	842
			993	0.006	0.987	620	1000
0.2	0.9	0.9	889	0.208	0.491	295	999
			976	0.003	0.677	621	1000
0.2	0.3	0.3	986	0.806	0.752	793	866
			1000	0.099	0.993	1000	1000
0.2	0.0	0.0	998	1.251	0.819	928	461
			1000	0.239	0.997	1000	1000
0.9	0.0	0.9	999	0.397	0.901	964	998
			1000	0.027	0.997	1000	1000
0.9	0.9	0.0	897	0.151	0.631	552	999
			978	0.001	0.685	612	1000
0.9	0.9	0.9	919	0.097	0.627	538	1000
			987	0.001	0.684	626	1000
0.9	0.3	0.3	985	0.313	0.826	849	999
			1000	0.016	0.993	1000	1000
0.9	0.0	0.0	997	0.551	0.910	867	979
			1000	0.032	0.997	1000	1000

From the numbers of $|t_{\hat{\gamma}}| \geq 1.96$, we confirm that in this case the nonsense regression phenomenon occurs and it is found that when $T \to \infty$ the $DW \to 0$ and $\overline{R}^2 \to 1$ as stated in Theorem 1.

Table 3.2: Results for DGPs (3), (4) where $\beta_x = 1$, $\mu_y = \mu_x = 0.5$.

β_y	ϕ_y	ϕ_x	$\lvert t_{\hat{\gamma}} \rvert \geq 1.96$	DW	$\overline{R}^2$	$\overline{R}^2 > 0.7$	$DW < 1.22$
0.2	0.0	0.9	910	0.823	0.568	460	753
			986	0.006	0.684	628	1000
0.2	0.9	0.0	995	0.460	0.678	557	996
			1000	0.089	0.995	1000	1000
0.2	0.9	0.9	874	0.297	0.469	294	998
			986	0.003	0.681	621	1000
0.2	0.3	0.3	980	0.939	0.749	783	766
			1000	0.103	0.993	1000	1000
0.2	0.0	0.0	997	1.461	0.824	934	278
			1000	0.240	0.997	1000	1000
0.9	0.0	0.9	905	0.146	0.623	526	995
			991	0.001	0.681	613	1000
0.9	0.9	0.0	1000	0.378	0.906	966	998
			1000	0.027	0.997	1000	1000
0.9	0.9	0.9	915	0.095	0.636	550	1000
			989	0.001	0.658	597	1000
0.9	0.3	0.3	991	0.295	0.834	857	1000
			1000	0.016	0.993	1000	1000
0.9	0.0	0.0	1000	0.543	0.912	961	971
			1000	0.033	0.997	1000	1000

From the numbers of $\lvert t_{\hat{\gamma}} \rvert \geq 1.96$, we confirm that in this case the nonsense regression phenomenon occurs and it is found that when $T \to \infty$ the $DW \to 0$ and $\overline{R}^2 \to 1$ as is shown in Theorem 2.

Table 3.3: Results for DGPs (5), (6) where $\beta_y = 0$, $\beta_x = 1$, $\mu_y = \mu_x = 0.5$.

ϕ_y	ϕ_x	$\lvert t_{\hat{\gamma}} \rvert \geq 1.96$	DW	$\overline{R}^2$	$\overline{R}^2 > 0.7$	$DW < 1.22$
0.0	0.9	62	2.051	0.000	0	2
		60	2.000	0.000	0	0
0.9	0.0	620	0.487	0.162	11	993
		653	0.212	0.017	0	1000
0.9	0.9	608	0.481	0.165	12	998
		666	0.212	0.018	0	1000
0.3	0.3	134	1.522	0.013	0	130
		144	1.402	0.001	0	1
0.0	0.0	48	2.072	-0.001	0	1000
		53	2.004	0.000	0	0

When $\phi_y = 0.0, DW \rightarrow 2$ and when $\phi_y = 0.9$, $DW \rightarrow 0.2$ because if $T \rightarrow \infty$, $DW \rightarrow 2\left(1 - \phi_y\right)$ as is shown in Theorem 3. The finite sample rejection times for the case $\phi_y = 0.0$ are 62, 60, 48, 53 respectively because if $\phi_y = 0.0$, $t_{\hat{\gamma}}$

converges in $N(0,1)$ as $T \rightarrow \infty$. In the cases for $\phi_y \neq 0.0$ we recognize that the nonsense regression phenomenon occurs.

Table 3.4: Results for DGPs (7), (8) where $\beta_y = 1$, $\beta_x = 0$, $\mu_y = \mu_x = 0.5$.

ϕ_y	ϕ_x	$\lvert t_{\hat{\gamma}} \rvert \geq 1.96$	DW	$\overline{R}^2$	$\overline{R}^2 > 0.7$	$DW < 1.22$
0.0	0.9	597	0.122	0.157	4	1000
		646	0.004	0.017	0	1000
0.9	0.0	47	0.061	0.000	0	1000
		49	0.002	0.000	0	1000
0.9	0.9	578	0.114	0.160	15	1000
		652	0.005	0.019	0	1000
0.3	0.3	149	0.086	0.014	0	1000
		163	0.003	0.001	0	1000
0.0	0.0	53	0.070	-0.001	0	1000
		55	0.002	0.000	0	1000

It is found that when $T \rightarrow \infty$ the $DW \rightarrow 0$ and $\overline{R}^2 \rightarrow 0$ as is explained by Theorem 4. The finite sample rejection times for the case $\phi_x = 0.0$ are 47,49,53,55 respectively because if $\phi_x = 0.0$, the t-statistics $t_{\hat{\gamma}}$ converges in $N(0,1)$ as $T \rightarrow \infty$. In the case of $\phi_x \neq 0.0$, we confirm that the nonsense regression phenomenon occurs.

4 Concluding Remarks

In this paper we considered the spurious or nonsense regression phenomenon where the DGPs of the regressor and the regressand were I(1) with drift vs. trend stationary and I(0) vs. I(1) with drift, and all of these patterns had a first order autoregressive errors.

In the former case, It is found that the nonsense regression phenomenon occurs. Comparing our Theorem 1 and 2 with the Theorem 1 of Kim, Lee and Newbold(2004), it is similar that $\hat{\gamma}$ converges in probability to the ratio of coefficients of the trend. But the asymptotic result of $t_{\hat{\gamma}}$ is slightly different. In our case it has a limiting distribution but in Kim, et al. case it has a probability limit. The DW does not converge in probability to 0 for the case of Theorem 1 of Kim, et al. and the R^2(or $\overline{R}^2$) converges to 1 faster than our cases.

On the other hand, in the latter case, it is found that the nonsense regression phenomenon also occurs. Our results imply that as the AR parameter ϕ approaches unity, the asymptotic rejection rate of the t-statistics $t_{\hat{\gamma}}$ becomes larger. Therefore, the phenomenon of the nonsense regression depends on closeness of ϕ to unity. Our Theorem 3 and 4 are similar with the Theorem 2 of Kim, et al. and the asymptotic behavior of the DW and R^2 are also similar.

Note: The detailed derivations of the Theorems are omitted but are obtainable from the author upon request.

References

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

Durlauf, S. N., and P.C.B. Phillips(1988) Trends versus random walks in time series analysis, *Econometrica*, 56, 1333-1354.

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

He Z., and K. Maekawa(2001) On spurious Granger causality. *Economics Letters*, 73, 307-313.

Kim,T.H., Lee, Y.S., and P. Newbold (2004) Spurious regressions with stationary processes around linear trends. *Economics Letters*, 83, 257-262.

Noriega, A. E, and D. V-Santaularia (2005) Spurious regression under broken-trend stationarity, *Journal of Time Series Analysis*, 27, 671-684.

Phillips, P.C.B. (1986) Understanding spurious regressions in econometrics. *Journal of Econometrics*, 33, 311-340.

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