UNIVERSITY OF NOTTINGHAM

Discussion Papers in Economics

Discussion Paper
No. 03/07

SPURIOUS REGRESSIONS WITH PROCESSES AROUND LINEAR TRENDS OR DRIFTS

by Tai-Hwan Kim, Young-Sook Lee and Paul Newbold

May 2003

DP 03/07
ISSN 1360-2438

UNIVERSITY OF NOTTINGHAM

Discussion Papers in Economics

Discussion Paper
No. 03/07

SPURIOUS REGRESSIONS WITH PROCESSES AROUND LINEAR TRENDS OR DRIFTS

by Tae-Hwan Kim, Young-Sook Lee and Paul Newbold

Tae-Hwan Kim is Lecturer, Young-Sook Lee is Research Fellow and Paul Newbold is Professor, all in the School of Economics, University of Nottingham

May 2003

Spurious Regressions with Processes around Linear Trends or Drifts

Tae-Hwan Kim, Young-Sook Lee and Paul Newbold

School of Economics, University of Nottingham, Nottingham NG7 2RD, UK

Summary In this paper we consider the situation where the deterministic components of the processes generating individual series are linear trends and the individual series are independent I(0) or I(1) processes. We show that when those time series are used in ordinary least squares regression, the phenomenon of spurious regression occurs regardless of whether a time trend is included in the regression.

1 Introduction

Following Granger and Newbold (1974) and Phillips (1986), the spurious regression phenomenon in econometrics is generally understood to refer to the impact of ordinary least squares regression on independent time series generated by integrated processes. However, as has recently been discussed in the literature, the phenomenon can also occur for regressions involving highly autocorrelated trend stationary processes.

In this paper we consider the situation where the deterministic components of the processes generating individual series are linear trends. In Section 2, the stochastic components are taken to be stationary first order autoregressions. We consider the cases where a time trend both is not and is included in the fitted regression. Section 3 of the paper extends these results to what might be viewed as the limiting case, where individual series are generated by random walks with drifts.

2 Spurious regressions with I(0) processes around linear trends

We consider two independent stationary processes y_t and x_t around trends, that are generated from the following DGP:

$$\begin{aligned} y_t &= \mu_y + \beta_y t + u_{yt} \quad ; \quad u_{yt} = \phi_y u_{yt-1} + \varepsilon_{yt} \quad ; \quad |\phi_y| < 1, \\ x_t &= \mu_x + \beta_x t + u_{xt} \quad ; \quad u_{xt} = \phi_x u_{xt-1} + \varepsilon_{xt} \quad ; \quad |\phi_x| < 1. \end{aligned} \tag{1}$$

Assumption 1. (i) ε_{yt} is i.i.d(0, σ_y^2), (ii) ε_{xt} is i.i.d(0, σ_x^2) and (iii) ε_{yt} and ε_{xt} are independent.

The conditions on ε_{yt} and ε_{xt} can be relaxed, but we keep Assumption1 for clarity and simplicity. To investigate a possible relationship between y_t and x_t, a researcher might run an OLS regression of y_t on a constant and x_t:

$$y_t = \hat{\alpha} + \hat{\gamma} x_t + e_t. \tag{2}$$

This regression has been extensively studied in the literature. When y_t and x_t are independent random walks without drifts, Granger and Newbold (1974) and Phillips (1986) analysed this regression showing that the OLS estimator $\hat{\gamma}$ converges to a random variable and its t-statistic diverges with a rate $T^{1/2}$. Examining the same regression, Entorf (1997) proved that if y_t and x_t are independent random walks with non-zero drifts, then $\hat{\gamma}$ converges in probability to the ratio of the drift of y_t and the drift of x_t, and its t-statistic diverges with the same speed $T^{1/2}$. Similar results were reported by Marmol (1998) in the context of nonstationary fractionally integrated processes. While most studies focus on independent nonstationary processes, Granger *et al* (2001) and Tsay and Chung (2000) studied the regression, assuming that y_t and x_t are independent stationary processes without any trend components.

Similarly to these previous studies, we examine the same regression as (2), but assuming that y_t and x_t are independent stationary processes around non-zero sloping trends; specifically at least one of β_y and β_x in (1) is not zero. The following theorems show the impact of these non-zero slope terms on the asymptotic distributions of various statistics from (2).

Theorem 1 *Suppose that y_t and x_t are generated by (1), the equation in (2) is estimated by least squares regression and Assumption 1 holds. If $\beta_y \neq 0$ and $\beta_x \neq 0$, then we have*

$$\hat{\gamma} \xrightarrow{p} \frac{\beta_y}{\beta_x},$$

$$T^{-3/2} t_{\hat{\gamma}} \xrightarrow{p} \frac{1}{\sqrt{12}} sgn(\beta_y) sgn(\beta_x) \left[\frac{\sigma_y^2}{\beta_y^2(1-\phi_y^2)} + \frac{\sigma_x^2}{\beta_x^2(1-\phi_x^2)} \right]^{-\frac{1}{2}}$$

where $sgn(z) = 1_{[z \geq 0]} - 1_{[z<0]}$.

As in Entorf (1997), the OLS estimator $\hat{\gamma}$ converges in probability to the ratio of the two trend coefficients, β_y/β_x, and its t-statistic $t_{\hat{\gamma}}$ diverges. However, $t_{\hat{\gamma}}$ diverges with a faster rate $T^{3/2}$ than Entorf (1997) where $t_{\hat{\gamma}}$ diverges with rate $T^{1/2}$. From the probability limit of $T^{-3/2} t_{\hat{\gamma}}$ in Theorem 1, it can be easily seen that (i) the farther away from zero either β_y or β_x is, the larger the absolute value of $t_{\hat{\gamma}}$ becomes, and (ii) the closer to one either ϕ_y or ϕ_x is, the smaller the absolute value of $t_{\hat{\gamma}}$ is. By way of illustration we graph the probability limit multiplied by $T^{3/2}$, which can predict the finite sample behavior of $t_{\hat{\gamma}}$. Figure 1(a) shows the probability limit multiplied by $T^{3/2}$ as a function of β_x with $T = 50$, $\beta_y = 0.03$, $\phi_y = \phi_x = 0.3$ and $\sigma_y = \sigma_x = 1$. Even with T as small as 50, the t-statistic becomes larger than 1.96 (the 5% critical value from N(0,1)) once β_x is greater than 0.03. The case with $T = 100$ and $\beta_y = 0.01$ is displayed in Figure 1(b). With β_y as low as 0.01, the approximated t-statistic is well above 1.96 for almost any values of β_x. When varying β_y while fixing β_x at the same set of values as in Figures 1(a) - 1(b), the same graphs are obtained, which is expected from the expression given in Theorem 1. Hence they are not reported. On the other hand, the impact of changing T is shown in Figure 1(c) with $\beta_y = \beta_x = 0.01$ and in Figure 1(d) with $\beta_y = \beta_x = 0.03$. When $\beta_y = \beta_x = 0.01$, the t-statistic becomes larger than 1.96 for $T \geq 80$. It is more evident when $\beta_y = \beta_x = 0.03$, in which case any values of $T \geq 30$ can cause spurious regressions.

We next consider the case in which only one of β_y, β_x is zero. The following theorem shows the asymptotics in this case.

Theorem 2 *Suppose that y_t and x_t are generated by (1), the equation in (2) is estimated by least squares regression and Assumption 1 holds. If $\beta_y = 0$ and $\beta_x \neq 0$, then we have*

$$T^{3/2}\hat{\gamma} \xrightarrow{d} N\left(0, \frac{12\sigma_y^2}{\beta_x^2(1-\phi_y)^2}\right) \quad and \quad t_{\hat{\gamma}} \xrightarrow{d} N\left(0, \frac{1-\phi_y^2}{(1-\phi_y)^2}\right).$$

If $\beta_y \neq 0$ and $\beta_x = 0$, then we have

$$T^{-1/2}\hat{\gamma} \xrightarrow{d} N\left(0, \frac{\beta_y^2(1-\phi_x^2)^2}{12\sigma_x^2\left(1-\phi_x\right)^2}\right) \quad and \quad t_{\hat{\gamma}} \xrightarrow{d} N\left(0, \frac{1-\phi_x^2}{(1-\phi_x)^2}\right).$$

The OLS estimator $\hat{\gamma}$ converges to zero when the dependent variable y_t has no trend and the independent variable x_t has one. On the other hand, the same OLS estimator diverges when the trend term is present only in the dependent variable. Even though the behavior of the two OLS estimators are completely different, their corresponding t-statistics converge to very similar normal distributions. In fact, the variances of the two limiting normal distributions are of the exact same functional form in which the argument is the AR parameter of the variable with no time trend term. For example, when $\beta_y = 0$, only ϕ_y determines the limiting distribution. Let ϕ be either ϕ_y or ϕ_x. Then it can be easily shown that (i) the variance term $\frac{1-\phi^2}{(1-\phi)^2}$ is greater than unity if and only if $0 < \phi < 1$ and (ii) the variance term is a monotonically increasing function of ϕ. The second property implies that as ϕ approaches unity, the asymptotic rejection rate becomes larger. This is a sharp contrast with the result in Theorem 1 where the t-statistic $t_{\hat{\gamma}}$ is a decreasing function of ϕ_y and ϕ_x. Therefore, for any stationary AR(1) processes, spurious regressions can occur and the extent of that phenomenon depends on the closeness of the AR parameter ϕ to unity. Given that we know the exact form of the variance of the limting normal distribution, it is possible to calculate the asymptotic rejection probability based on $t_{\hat{\gamma}}$ and the 5% critical value (1.96) from N(0,1). Let Z_ϕ denote the limiting normal random variable. Then, the asymptotic rejection probability, denoted $ARP(\phi)$, is given by

$$\begin{aligned} ARP(\phi) &= \Pr(|Z_\phi| > 1.96) \\ &= 2\left[1 - \Phi(1.96v(\phi)^{-1/2})\right] \end{aligned} \tag{3}$$

where Φ is the cumulative distribution function of N(0,1) and $v(\phi) = \frac{1-\phi^2}{(1-\phi)^2}$. Figure 2 shows the graph of $ARP(\phi)$ against ϕ. It increases at a faster rate as ϕ approaches one.

All the results obtained so far are of interest when the trend components β_y or β_x are small so that the researcher is mistaken in believing that the regression in (2) is correctly specified. Of course, when the trend components are large enough to be easily detected, then a time trend term can be added to the regression;

$$y_t = \hat{\alpha} + \hat{\beta} t + \hat{\gamma} x_t + e_t. \tag{4}$$

However, the following theorem shows that adding a time trend does not eliminate spurious effects.

Theorem 3 *Suppose that y_t and x_t are generated by (1), the equation in (4) is estimated by least squares regression and Assumption 1 holds. Then*

$$T^{1/2}\hat{\gamma} \xrightarrow{d} N\left(0, \frac{(1-\phi_x^2\phi_y^2)\sigma_y^2}{(1-\phi_x^2)(1-\phi_x\phi_y)^2\sigma_x^2}\right),$$

$$t_{\hat{\gamma}} \xrightarrow{d} N\left(0, \frac{1-\phi_x^2\phi_y^2}{(1-\phi_x\phi_y)^2}\right).$$

In contrast with Theorems 1 and 2 in which different asymptotics result from different values of β_y and β_x, the results in Theorem 1 are invariant to β_y and β_x; they can be either zero or non-zero without affecting the limiting distributions. The t-statistic $t_{\hat{\gamma}}$ does converge to a normal distribution, but the variance of that distribution is not unity, but a complicated function of the AR parameters, ϕ_y and ϕ_x, so that unless $\phi_y\phi_x = 0$, using critical values from N(0,1) can lead to spurious rejections. Note that the limiting distribution of $t_{\hat{\gamma}}$ is the same as the one in Granger *et al* (2001) in which it is assumed that $\beta_y = \beta_x = 0$ and no trend term is included in the regression. Hence, the results can be regarded as an extension of Granger *et al* (2001).

In order to investigate the finite sample properties of the t-statistic $t_{\hat{\gamma}}$ in Theorems 1-3, we conduct some Monte Carlo simulations. The two error terms ε_{yt} and ε_{xt} are drawn from N(0,1) and various values of T are used: $T = 100, 500, 1000, 2000$ and 10000. The results based on regressions in (2) and (4) are given in Table 1 and Table 2 respectively. For comparison we also include different types of stationary DGPs: MA(1) and I(d) with $d \in (-0.5, 0.5)$.

We first discuss Table 1. The case of $\beta_y = \beta_x = 0$ has already been reported in the literature, but we include the case here for comparison purpose. When the error terms follow AR(1) processes and both trend components are non-zero, finite sample rejection rates are 100% regardless the values of other model parameters. This can be easily explained by Theorem 1. For example, for $\phi_y = \phi_x = 0.9, T = 100$ and $\beta_y = \beta_x = 0.2$, the calculation of the probability limit of $T^{-3/2}t_{\hat{\gamma}}$ in Theorem 1 predicts that $t_{\hat{\gamma}}$ is $100^{3/2} \times 0.0178 = 17.8$ which is well above 1.96. As T increases, or either ϕ_y or ϕ_x decreases, the magnitude of $t_{\hat{\gamma}}$ increases even further. Hence, the rejection rate of 100% is well predicted. When either β_y or β_x is zero, Theorem 2 predicts that the distribution of $t_{\hat{\gamma}}$ depends only on the AR parameter of the variable with no time trend. The finite sample rejection rates for the case of $\beta_y = 0, \beta_x = 0.2$ and $T = 100$, are 5.9%, 15.9% and 67% for $\phi_y =$ 0, 0.3, and 0.9 respectively (when $\phi_x = \phi_y$) and 4.8%, 68.2% for $\phi_y = 0$, 0.9 respectively (with $\phi_x = 0.9, 0$ respectively). The asymptotic rejection probability in (3) predicts that $ARP(\phi_y = 0) = 5\%$, $ARP(\phi_y = 0.3) =15\%$ and $ARP(\phi_y = 0.9) = 65\%$. As is obvious from the numbers, the finite sample rejection rates are very well explained by the asymptotic rejection probabilities whether ϕ_x is equal to ϕ_y or not.

In contrast to the AR(1) cases, when the error terms are generated by MA(1) processes, there are no spurious effects if either β_y or β_x is zero, in which case finite sample rejection rates are virtually zero. When $\beta_y = \beta_x =$ 0.2, the rejection rate is 100% as in the AR(1) cases. When the error terms are fractionally integrated with $d \in (-0.5, 0.5)$, the phenomenon of spurious regressions is evident. Tsay and Chung (2000) showed that spurious effects occur if the sum of the long memory parameters for y_t and x_t is greater than 0.5. When $d = 0.4$ for both y_t and x_t, the sum is 0.8 and rejection rates range from 20% to 100% depending on the values of β_y and β_x. Hence, the spurious effects in the case of $d = 0.4$ confirm the findings of Tsay and Chung (2000). However, it can be seen that spurious effects can also occur (see the case of $d = 0.1$) even when the sum of the two long memory parameters is less than 0.5. This finding has not been reported in the literature.

Table 2 shows finite sample rejection rates when a linear time trend term is added to the regression. First of all, all rejection rates for AR(1) cases become smaller than the corresponding ones in Table 1. As predicted by Theorem 3, the rejection rates are independent of the values of β_y or β_x. We also note that when either ϕ_y or ϕ_x is zero, there is no spurious rejection phenomenon. This is not caused by the asymmetry of β_y or β_x. Rather, it is simply because the asymptotic variance of $t_{\hat{\gamma}}$ is $\frac{1-\phi_x^2\phi_y^2}{(1-\phi_x\phi_y)^2}$ so that the variance becomes unity when either ϕ_y or ϕ_x is zero. It is interesting to see that while adding a time trend term reduces spurious effects in all AR(1) cases, the same thing produces spurious effects in the MA(1) cases; the rejection rates are about 7% for $\theta_y = \theta_x = 0.3$ and about 11% for $\theta_y = \theta_x = 1$.

3 Spurious regressions with I(1) processes with drifts

In this section we consider two independent random walks with drifts as in Entorf (1997):

$$\begin{aligned} y_t &= \mu_y + \beta_y t + u_{yt} \quad ; \quad u_{yt} = u_{yt-1} + \varepsilon_{yt}, \\ x_t &= \mu_x + \beta_x t + u_{xt} \quad ; \quad u_{xt} = u_{xt-1} + \varepsilon_{xt}. \end{aligned} \tag{5}$$

While Entorf (1997) examines the spurious regression phenomenon arising when the regression in (2) is employed, we study the issue using the regression in (4); that is we add a time trend term to his regression equation. The following theorem shows the asymptotic distributions of some statistics from the regression in (4).

Theorem 4 *Suppose that y_t and x_t are generated by (5), the equation in (4) is estimated by least squares regression and Assumption 1 holds. Then we have*

$$\begin{aligned}
T^{-1/2}\hat{\alpha} &\Rightarrow \frac{2\sigma_y(2H_yG_x - 6H_yM_x^2 - 3M_yG_x + 6M_yH_xM_x + 3P_{xy}M_x - 2P_{xy}H_x)}{G_x - 12M_x^2 + 12H_xM_x - 4H_x^2}, \\
\hat{\gamma} &\Rightarrow \frac{\sigma_y(6H_yM_x - 4H_yH_x - 12M_yM_x + 6M_yH_x + P_{xy})}{\sigma_x\left(G_x - 12M_x^2 + 12H_xM_x - 4H_x^2\right)} \equiv \Psi, \\
\hat{\beta} &\Rightarrow \beta_y - \Psi\beta_x, \\
R^2 &= O_p(1), \\
DW &= O_p(T^{-1}), \\
t_{\hat{\alpha}} &= O_p(T^{1/2}), \\
t_{\hat{\beta}} &= O_p(T^{1/2}), \\
t_{\hat{\gamma}} &= O_p(T^{1/2}),
\end{aligned}$$

where

$$\begin{aligned}
H_x &\equiv \int_0^1 W(r)dr, \quad H_y \equiv \int_0^1 V(r)dr, \\
M_x &\equiv \int_0^1 rW(r)dr, \quad M_y \equiv \int_0^1 rV(r)dr, \\
G_x &\equiv \int_0^1 W(r)^2dr, \quad P_{xy} \equiv \int_0^1 W(r)V(r)dr.
\end{aligned}$$

Here, $V(r)$ and $W(r)$ are independent Brownian motion processes defined as

$$\begin{aligned}
\sigma_y^{-1}T^{-1/2}\sum_{t=1}^{sT}\varepsilon_{yt} &\Rightarrow V(s), \\
\sigma_x^{-1}T^{-1/2}\sum_{t=1}^{sT}\varepsilon_{xt} &\Rightarrow W(s).
\end{aligned}$$

The results in Theorem 4 is an extension of Phillips (1986); specifically the OLS estimator $\hat{\gamma}$ converges to a random variable and its t-statistic $t_{\hat{\gamma}}$ diverges with rate $T^{1/2}$. This should not be surprising since the regression in (4) is identical to the following regression

$$\tilde{y}_t = \tilde{\alpha} + \hat{\gamma}\tilde{x}_t + e_t \tag{6}$$

where $\tilde{\alpha} = 0$, and $\tilde{y}_t$ and $\tilde{x}_t$ are the residuals from the regressions of y_t and x_t on $[1, t]$ respectively. Hence, the only difference between our results and Phillips (1986) is that in the above regression (6) $\tilde{y}_t$ and $\tilde{x}_t$ are pure random walks in Phillips (1986) and demeaned and detrended random walks in our case. Tables 3 and 4 show the spurious phenomenon in finite samples.

References

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

[2] Granger, C.W.J., Hyung, N. and Jeon, Y. (2001). Spurious regressions with stationary series, Applied Economics, 33, 899-904.

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

[4] Hassler, Uwe. (1996). Spurious regressions when stationary regressors are included. Economics Letters, 50, 25-31.

[5] Marmol, F. (1998). Spurious regression theory with nonstationary fractionally integrated processes. Journal of Econometrics, 84, 233-250.

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

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

4 Appendix

4.1 Proof of Theorem 1

The OLS estimator $\hat{\gamma}$ is given by $\hat{\gamma} = \left\{\sum_{t=1}^{T} (x_t - \bar{x})^2\right\}^{-1} \left\{\sum_{t=1}^{T} (x_t - \bar{x})(y_t - \bar{y})\right\}$. When $\beta_y \neq 0$ and $\beta_x \neq 0$, we have the following results: $T^{-3} \sum x_t^2 \xrightarrow{p} \beta_x^2/3$, $T^{-2} \sum x_t \xrightarrow{p} \beta_x/2, T^{-2} \sum y_t \xrightarrow{p} \beta_y/2$ and $T^{-3} \sum_{t=1}^{T} x_t y_t \xrightarrow{p} \beta_x \beta_y/3$. By combining these results, one can show that $T^{-3} \sum_{t=1}^{T} (x_t - \bar{x})^2 \xrightarrow{p} \beta_x^2/12$ and $T^{-3} \sum_{t=1}^{T} (x_t - \bar{x})(y_t - \bar{y}) \xrightarrow{p} \beta_x \beta_y/12$, which implies that $\hat{\gamma} \xrightarrow{p} \beta_y/\beta_x$. The t-statistic $t_{\hat{\gamma}}$ scaled by $T^{-3/2}$ is given by $T^{-3/2} t_{\hat{\gamma}} = (\hat{\sigma}^2)^{-1/2} \left\{T^{-3} \sum_{t=1}^{T} (x_t - \bar{x})^2\right\}^{1/2} \hat{\gamma}$

where $\hat{\sigma}^2 = T^{-1}\sum_{t=1}^{T} e_t^2$. It is straightforward to show that $\hat{\sigma}^2 \xrightarrow{p} \sigma_y^2(1-\phi_y^2)^{-1} + (\beta_y/\beta_x)^2\sigma_{x^2}(1-\phi_x^2)^{-1}$. Hence, we have

$$T^{-3/2}\, t_{\hat{\gamma}} \xrightarrow{p} \left\{ \frac{\sigma_y^2}{1-\phi_y^2} + \left(\frac{\beta_y}{\beta_x}\right)^2 \frac{\sigma_x^2}{1-\phi_x^2} \right\}^{-\frac{1}{2}} \left(\frac{\beta_x^2}{12}\right)^{1/2} \frac{\beta_y}{\beta_x}$$

which simplifies to the expression in the theorem.

4.2 Proof of Theorem 2

We first consider the case where $\beta_y = 0$ and $\beta_x \neq 0$. Since $\beta_x \neq 0$, we still have $T^{-3}\sum_{t=1}^{T}(x_t - \bar{x})^2 \xrightarrow{p} \beta_x^2/12$. The numerator of $\hat{\gamma}$ scaled by $T^{-3/2}$ is given by

$$\begin{aligned} T^{-3/2}\sum_{t=1}^{T}(x_t-\bar{x})(y_t-\bar{y}) &= \beta_x A' \begin{pmatrix} T^{-1/2}\sum_{t=1}^{T} u_{yt} \\ T^{-3/2}\sum_{t=1}^{T} t u_{yt} \end{pmatrix} + o_p(1) \\ &\xrightarrow{d} \beta_x A' N\left(0, \frac{\sigma_y^2}{(1-\phi_y)^2} Q\right) \\ &= N\left(0, \frac{\beta_x^2\sigma_y^2}{12(1-\phi_y)^2}\right) \end{aligned}$$

where $A = (-1/2, 1)'$ and $Q = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{bmatrix}$. Hence, we have $T^{3/2}\hat{\gamma} \xrightarrow{d} N\left(0, \frac{12\sigma_y^2}{\beta_x^2(1-\phi_y)^2}\right)$. Next we turn to the t-statistic $t_{\hat{\gamma}}$ which is given by $t_{\hat{\gamma}} = (\hat{\sigma}^2)^{-1/2}\left\{T^{-3}\sum_{t=1}^{T}(x_t-\bar{x})^2\right\}^{1/2} T^{3/2}\hat{\gamma}$. Using the fact that $\hat{\sigma}^2 \xrightarrow{p} \sigma_y^2(1-\phi_y^2)^{-1}$, it can be shown that

$$t_{\hat{\gamma}} \xrightarrow{d} \left(\frac{\sigma_y^2}{1-\phi_y^2}\right)^{-1/2} \left(\frac{\beta_x^2}{12}\right)^{1/2} N\left(0, \frac{12\sigma_y^2}{\beta_x^2(1-\phi_y)^2}\right) = N\left(0, \frac{1-\phi_y^2}{(1-\phi_y)^2}\right).$$

Now we consider the next case of $\beta_y \neq 0$ and $\beta_x = 0$. Since $\beta_x = 0$, we have $T^{-1}\sum_{t=1}^{T}(x_t-\bar{x})^2 \xrightarrow{p} \sigma_x^2\left(1-\phi_x^2\right)^{-1}$ by a law of large numbers. Similarly to the previous case, we have

$$\begin{aligned} T^{-3/2}\sum_{t=1}^{T}(x_t-\bar{x})(y_t-\bar{y}) &= \beta_y A' \begin{pmatrix} T^{-1/2}\sum_{t=1}^{T} u_{xt} \\ T^{-3/2}\sum_{t=1}^{T} t u_{xt} \end{pmatrix} + o_p(1) \\ &\xrightarrow{d} \beta_y A' N\left(0, \frac{\sigma_x^2}{(1-\phi_x)^2} Q\right) \\ &= N\left(0, \frac{\beta_y^2\sigma_x^2}{12(1-\phi_x)^2}\right) \end{aligned}$$

By combining the results, we have $T^{-1/2}\hat{\gamma} \xrightarrow{d} N\left(0, \frac{\beta_y^2(1-\phi_x^2)^2}{12\sigma_x^2(1-\phi_x)^2}\right)$. The t-statistic is given by $t_{\hat{\gamma}} = (T^{-2}\hat{\sigma}^2)^{-1/2}\left\{T^{-1}\sum_{t=1}^{T}(x_t-\bar{x})^2\right\}^{1/2} T^{1/2}\hat{\gamma}$ for which we have

$$t_{\hat{\gamma}} \xrightarrow{d} \left(\frac{12}{\beta_y^2}\right)^{1/2}\left(\frac{\sigma_x^2}{1-\phi_x^2}\right)^{1/2} N\left(0, \frac{\beta_y^2\sigma_x^2}{12(1-\phi_x)^2}\right) = N\left(0, \frac{1-\phi_x^2}{(1-\phi_x)^2}\right)$$

using $T^{-2}\hat{\sigma}^2 \xrightarrow{p} \beta_y^2/12$. This completes the proof.

4.3 Proof of Theorem 3

We can rewrite $y_t = \hat{\alpha} + \hat{\beta}t + \hat{\gamma}x_t + e_t$ as $u_{yt} = \hat{\alpha} + (\hat{\beta} - \beta_y + \hat{\gamma}\beta_x)t + \hat{\gamma}u_{xt} + e_t$. Then, we have $\hat{\delta} = (\sum_{t=1}^{T} Z_t Z_t')^{-1}\sum_{t=1}^{T} Z_t u_{yt}$ where $Z_t = (1,\ t,\ u_{xt})'$ and $\hat{\delta} = (\hat{\alpha},\ \hat{\beta} - \beta_y + \hat{\gamma}\beta_x,\ \hat{\gamma})'$. We define $D_T = diag(T^{1/2}, T^{3/2}, T^{1/2})$ and pre-multiply both sides of the expression of $\hat{\delta}$ by D_T to obtain

$$\begin{bmatrix} T^{1/2}\hat{\alpha} \\ T^{3/2}(\hat{\beta} - \beta_y + \hat{\gamma}\beta_x) \\ T^{1/2}\hat{\gamma} \end{bmatrix} = \left(D_T^{-1}\sum_{t=1}^{T} Z_t Z_t' D_T^{-1}\right)^{-1} D_T^{-1}\sum_{t=1}^{T} Z_t u_{yt}.$$

It can be shown that

$$D_T^{-1}\left(\sum_{t=1}^{T} Z_t Z_t'\right) D_T^{-1} \xrightarrow{p} \begin{bmatrix} 1 & 1/2 & 0 \\ 1/2 & 1/3 & 0 \\ 0 & 0 & \sigma_x^2/(1-\phi_x^2) \end{bmatrix}.$$

Using the fact that the off-diagonal terms are zero, we have

$$\begin{aligned} T^{1/2}\hat{\gamma} &= \frac{1-\phi_x^2}{\sigma_x^2} T^{-1/2}\sum_{t=1}^{T} u_{xt}u_{yt} + o_p(1) \\ &\xrightarrow{d} \frac{1-\phi_x^2}{\sigma_x^2} N\left(0, \frac{1-\phi_x^2\phi_y^2}{(1-\phi_x^2)(1-\phi_y^2)(1-\phi_x^2\phi_y^2)}\ \sigma_x^2\sigma_y^2\right) \\ &= N\left(0, \frac{(1-\phi_x^2)(1-\phi_x^2\phi_y^2)\sigma_y^2}{(1-\phi_y^2)(1-\phi_x\phi_y)\sigma_x^2}\right). \end{aligned}$$

The t-statistic $t_{\hat{\gamma}}$ is given by $t_{\hat{\gamma}} = \hat{\gamma}\left\{v\hat{a}r(\hat{\gamma})\right\}^{-1/2}$ where $v\hat{a}r(\hat{\gamma}) = \hat{\sigma}^2 i_3'\left(\sum_{t=1}^{T} Z_t Z_t'\right)^{-1} i_3$, $\hat{\sigma}^2 = T^{-1}\sum_{t=1}^{T} e_t^2$ and $i_3 = (0\ 0\ 1)'$. Note that

$$\begin{aligned} \hat{\sigma}^2 &= T^{-1}\sum_{t=1}^{T}\left\{(u_{yt} - \bar{u}_y) - \left(\hat{\beta} - \beta_y + \hat{\gamma}\beta_x\right)(t - \bar{t}) - \hat{\gamma}\,(u_{xt} - \bar{u}_x)\right\}^2 \\ &\xrightarrow{p} \frac{\sigma_y^2}{1-\phi_y^2} \end{aligned}$$

which implies that

$$t_{\hat{\gamma}} \xrightarrow{d} N\left(0, \frac{1-\phi_x^2\phi_y^2}{(1-\phi_x\phi_y)^2}\right).$$

4.4 Proof of Theorem 4

We can rewrite $y_t = \hat{\alpha} + \hat{\beta}t + \hat{\gamma}x_t + e_t$ as

$$u_{yt} = \hat{\alpha} + \tilde{\beta}t + \hat{\gamma}u_{xt} + e_t \tag{7}$$

where $\tilde{\beta} = \hat{\beta} - \beta_y + \hat{\gamma}\beta_x$. Then, we have $\hat{\delta} = (\sum_{t=1}^T Z_t Z_t')^{-1} \sum_{t=1}^T Z_t u_{yt}$ where $Z_t = (1,\ t,\ u_{xt})'$ and $\hat{\delta} = (\hat{\alpha}, \tilde{\beta}, \hat{\gamma})'$; that is,

$$\hat{\delta} = \begin{bmatrix} T & \sum_{t=1}^T t & \sum_{t=1}^T u_{xt} \\ \sum_{t=1}^T t & \sum_{t=1}^T t^2 & \sum_{t=1}^T t u_{xt} \\ \sum_{t=1}^T u_{xt} & \sum_{t=1}^T t u_{xt} & \sum_{t=1}^T u_{xt}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^T u_{yt} \\ \sum_{t=1}^T t u_{yt} \\ \sum_{t=1}^T u_{xt} u_{yt} \end{bmatrix} \tag{8}$$

The limits of individual terms in the above expression when properly scaled are as follows:

$$\begin{aligned}
T^{-3/2}\sum_{t=1}^T u_{xt} &\Rightarrow \sigma_x \int_0^1 W(r)dr \equiv \sigma_x H_x, \\
T^{-5/2}\sum_{t=1}^T t u_{xt} &\Rightarrow \sigma_x \int_0^1 rW(r)dr \equiv \sigma_x M_x, \\
T^{-2}\sum_{t=1}^T u_{xt}^2 &\Rightarrow \sigma_x^2 \int_0^1 W(r)^2 dr \equiv \sigma_x G_x, \\
T^{-3/2}\sum_{t=1}^T u_{yt} &\Rightarrow \sigma_y \int_0^1 V(r)dr \equiv \sigma_y H_y, \\
T^{-5/2}\sum_{t=1}^T t u_{yt} &\Rightarrow \sigma_y \int_0^1 rV(r)dr \equiv \sigma_y M_y, \\
T^{-2}\sum_{t=1}^T u_{xt}u_{yt} &\Rightarrow \sigma_x\sigma_y \int_0^1 W(r)V(r)dr \equiv \sigma_x\sigma_y P_{xy}.
\end{aligned}$$

Pre-multiplying both sides of (8) by $diag(T^{-1/2}, T^{1/2}, 1)$ and combining the above results, we have

$$\begin{aligned}
T^{-1/2}\hat{\alpha} &\Rightarrow \frac{2\sigma_y(2H_yG_x - 6H_yM_x^2 - 3M_yG_x + 6M_yH_xM_x + 3P_{xy}M_x - 2P_{xy}H_x)}{G_x - 12M_x^2 + 12H_xM_x - 4H_x^2}, \\
\hat{\gamma} &\Rightarrow \frac{\sigma_y(6H_yM_x - 4H_yH_x - 12M_yM_x + 6M_yH_x + P_{xy})}{\sigma_x\left(G_x - 12M_x^2 + 12H_xM_x - 4H_x^2\right)} \equiv \Psi, \\
\hat{\beta} &\Rightarrow \beta_y - \Psi\beta_x.
\end{aligned}$$

Noting that the two regressions in (4) and (7) are equivalent, the coefficient of determination R^2 is given by

$$\begin{aligned} R^2 &= 1 - \frac{\sum_{t=1}^T e_t^2}{\sum_{t=1}^T (u_{yt} - \bar{u}_y)^2} = 1 - \frac{\sum_{t=1}^T e_t^2}{\sum_{t=1}^T u_{yt}^2 - T\bar{u}_y^2} \\ &= 1 - \frac{O_p(T^2)}{O_p(T^2)}. \end{aligned}$$

Therefore, $R^2 = O_p(1)$. The Durbin-Watson statistic DW is

$$\begin{aligned} DW &= \frac{\sum_{t=1}^T (e_t - e_{t-1})^2}{\sum_{t=1}^T e_t^2} \\ &= \frac{\sum_{t=1}^T \{\varepsilon_{yt} - \tilde{\beta} - \hat{\gamma}\varepsilon_{xt}\}^2}{\sum_{t=1}^T e_t^2} \\ &= \frac{O_p(T^1)}{O_p(T^2)}. \end{aligned}$$

Therefore, $DW = O_p(T^{-1})$. The t-statistic $t_{\hat{\gamma}}$ is $t_{\hat{\gamma}} = \hat{\gamma}\{v\hat{a}r(\hat{\gamma})\}^{-1/2}$ where $v\hat{a}r(\hat{\gamma}) = \hat{\sigma}^2 i_3' \left(\sum_{t=1}^T Z_t Z_t'\right)^{-1} i_3$, $\hat{\sigma}^2 = T^{-1}\sum_{t=1}^T e_t^2$ and $i_3 = (0\ 0\ 1)'$. Since $\hat{\sigma}^2 = T^{-1}\sum_{t=1}^T e_t^2 = O_p(T)$, $v\hat{a}r(\hat{\gamma}) = O_p(T^{-1/2})$ and $\hat{\gamma} = O_p(1)$, we have $t_{\hat{\gamma}} = O_p(T^{1/2})$. The proofs for the other claims in the theorem are obtained in a similar way and hence they are omitted.

Figure 1(a). $T^{3/2}$Plim with T = 50, β_y = 0.03

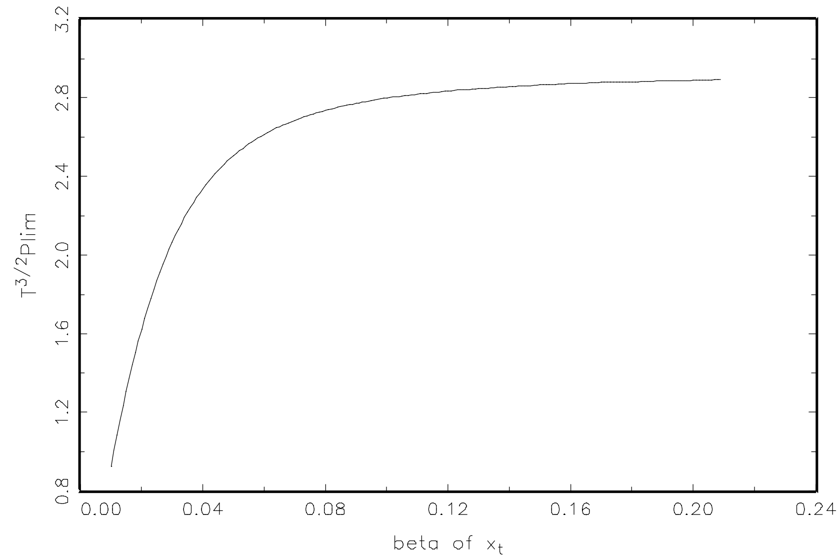


Figure 1(b). $T^{3/2}$Plim with T = 100, β_y = 0.01

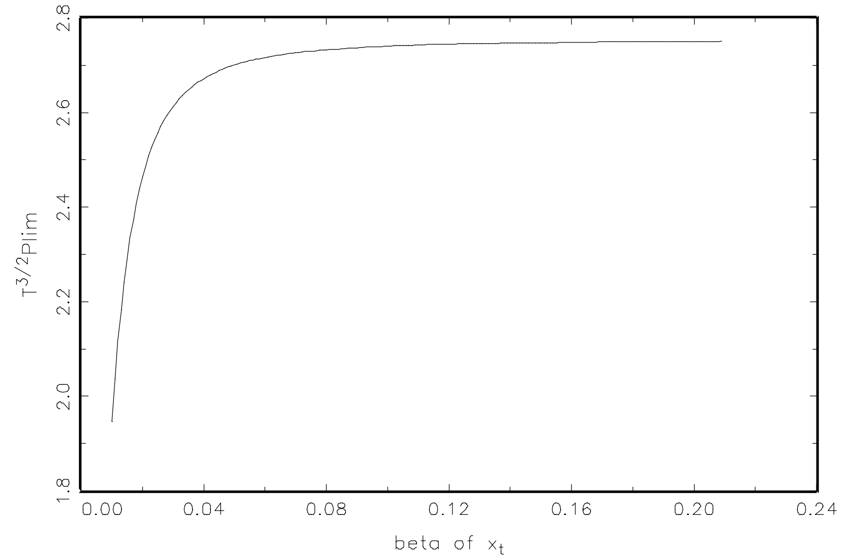


Figure 1(c). $T^{3/2}$Plim with β_y = β_x = 0.01

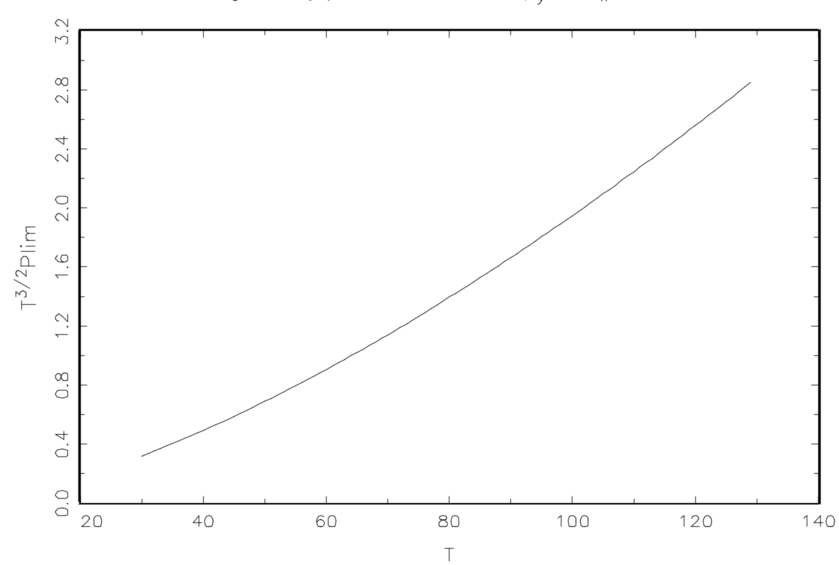
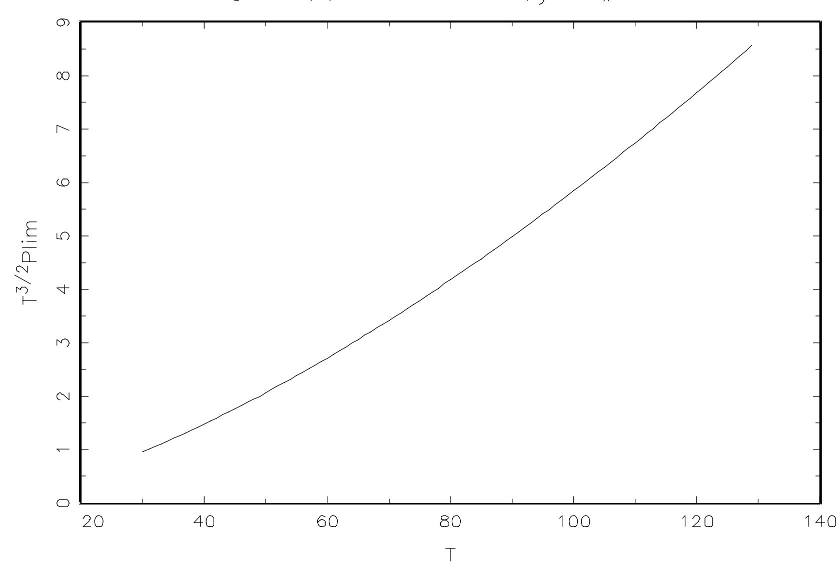


Figure 1(d). $T^{3/2}$Plim with β_y = β_x = 0.03



GAUSS   Wed Dec 18 12:23:35 2002

Figure 2. Rejection Probability

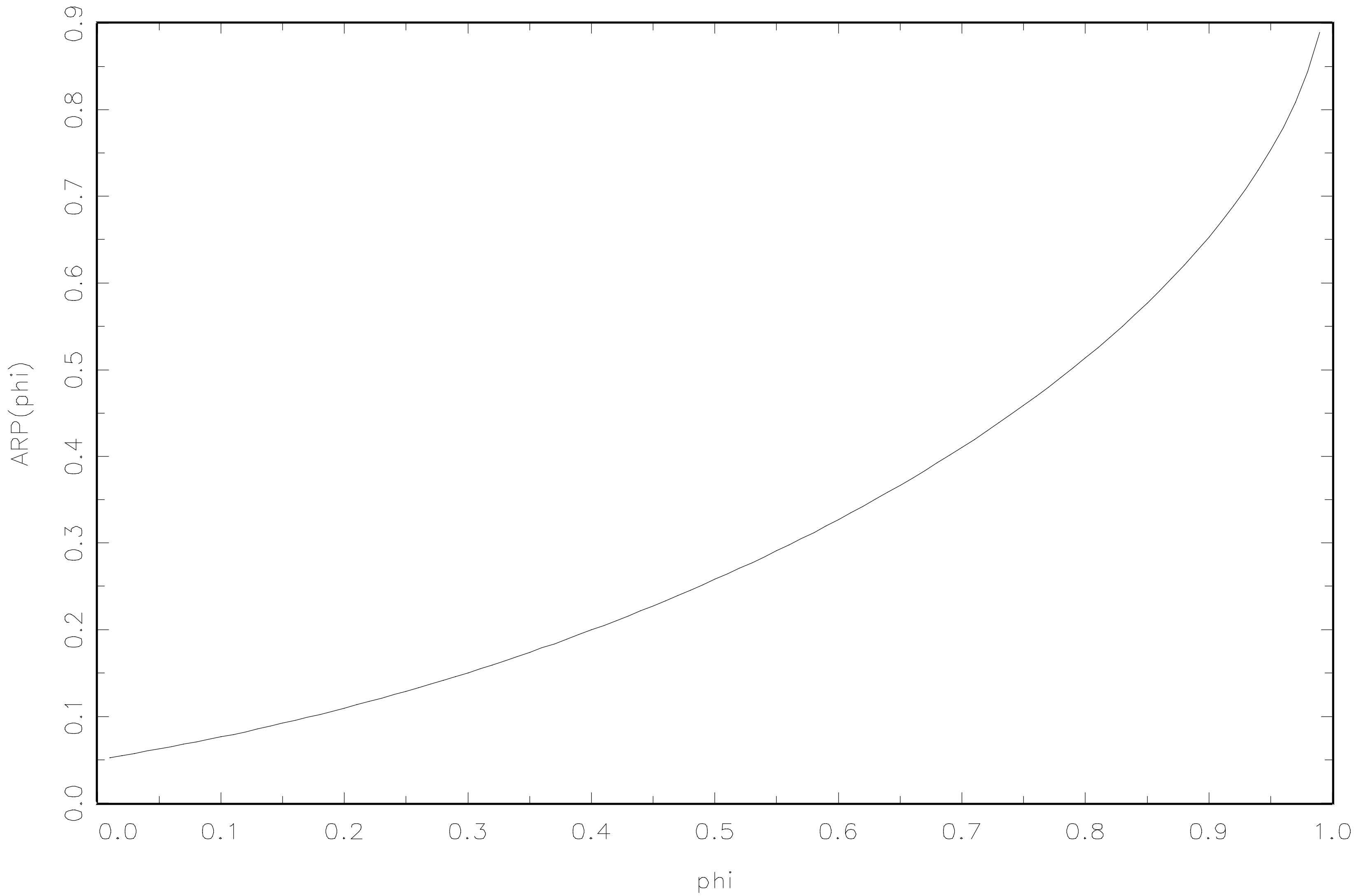


Table 1. Proportion of rejections ($|t_{\hat{\gamma}}| > 1.96$) based on the regression $y_t = \hat{\alpha} + \hat{\gamma} x_t + e_t$ when y_t and x_t are independent stationary processes around linear trends.

T	β_y	β_x	AR(1) $\phi_y = 0$ $\phi_x = 0$	AR(1) $\phi_y = 0.3$ $\phi_x = 0.3$	AR(1) $\phi_y = 0.9$ $\phi_x = 0.9$	AR(1) $\phi_y = 0$ $\phi_x = 0.9$	AR(1) $\phi_y = 0.9$ $\phi_x = 0$	MA(1) $\theta_y = 0.3$ $\theta_x = 0.3$	MA(1) $\theta_y = 1$ $\theta_x = 1$	I(d) $d = 0.1$	I(d) $d = 0.4$
50	0	0	0.052	0.073	0.478	0.059	0.053	0.073	0.112	0.065	0.138
	0	0.2	0.053	0.145	0.617	0.066	0.662	0.009	0.001	0.119	0.444
	0.2	0	0.065	0.150	0.614	0.652	0.048	0.013	0	0.126	0.420
	0.2	0.2	1.000	1.000	0.962	0.988	0.995	1.000	1.000	1.000	1.000
100	0	0	0.052	0.074	0.492	0.048	0.047	0.079	0.111	0.064	0.209
	0	0.2	0.059	0.159	0.670	0.048	0.682	0.003	0	0.153	0.551
	0.2	0	0.059	0.160	0.652	0.656	0.067	0.005	0	0.137	0.499
	0.2	0.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
500	0	0	0.041	0.065	0.523	0.040	0.043	0.058	0.096	0.068	0.379
	0	0.2	0.050	0.162	0.671	0.049	0.668	0.005	0	0.212	0.737
	0.2	0	0.047	0.145	0.670	0.666	0.047	0.009	0	0.224	0.757
	0.2	0.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2000	0	0	0.055	0.070	0.544	0.044	0.050	0.067	0.091	0.064	0.530
	0	0.2	0.043	0.152	0.640	0.044	0.639	0.003	0	0.257	0.830
	0.2	0	0.049	0.158	0.631	0.631	0.048	0.004	0	0.259	0.817
	0.2	0.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
10000	0	0	0.054	0.087	0.544	0.051	0.054	0.007	0.123	0.060	0.583
	0	0.2	0.053	0.149	0.683	0.053	0.683	0.006	0	0.264	0.833
	0.2	0	0.036	0.129	0.656	0.654	0.036	0.002	0	0.263	0.834
	0.2	0.2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Notes: 1. MA(1) DGP: $y_t = \mu_y + \beta_y t + u_{yt};\ u_{yt} = \varepsilon_{yt} - \theta_y \varepsilon_{yt-1}$,
$x_t = \mu_x + \beta_x t + u_{xt};\ u_{xt} = \varepsilon_{xt} - \theta_x \varepsilon_{xt-1}$.

2. I(d) DGP: $y_t = \mu_y + \beta_y t + u_{yt};\ (1-L)^d u_{yt} = \varepsilon_{yt}$,
$x_t = \mu_x + \beta_x t + u_{xt};\ (1-L)^d u_{xt} = \varepsilon_{xt}$.

Table 2. Proportion of rejections ($|t_{\hat{\gamma}}| > 1.96$) based on the regression $y_t = \hat{\alpha} + \hat{\beta}t + \hat{\gamma}x_t + e_t$ when y_t and x_t are independent stationary processes around linear trends.

T	β_y	β_x	AR(1) $\phi_y = 0$ $\phi_x = 0$	AR(1) $\phi_y = 0.3$ $\phi_x = 0.3$	AR(1) $\phi_y = 0.9$ $\phi_x = 0.9$	AR(1) $\phi_y = 0$ $\phi_x = 0.9$	AR(1) $\phi_y = 0.9$ $\phi_x = 0$	MA(1) $\theta_y = 0.3$ $\theta_x = 0.3$	MA(1) $\theta_y = 1$ $\theta_x = 1$	I(d) $d = 0.1$	I(d) $d = 0.4$
50	0	0	0.052	0.080	0.432	0.056	0.042	0.071	0.110	0.073	0.121
	0	0.2	0.062	0.092	0.423	0.061	0.049	0.069	0.124	0.069	0.135
	0.2	0	0.066	0.072	0.436	0.059	0.043	0.077	0.111	0.053	0.117
	0.2	0.2	0.059	0.084	0.433	0.058	0.058	0.063	0.101	0.056	0.103
100	0	0	0.053	0.085	0.499	0.050	0.054	0.077	0.110	0.066	0.170
	0	0.2	0.055	0.074	0.479	0.053	0.056	0.069	0.111	0.058	0.162
	0.2	0	0.059	0.078	0.493	0.056	0.053	0.071	0.118	0.037	0.163
	0.2	0.2	0.045	0.064	0.457	0.048	0.046	0.071	0.105	0.054	0.151
500	0	0	0.042	0.060	0.520	0.043	0.039	0.057	0.096	0.064	0.317
	0	0.2	0.044	0.081	0.502	0.051	0.048	0.073	0.122	0.054	0.276
	0.2	0	0.062	0.080	0.510	0.050	0.050	0.065	0.107	0.062	0.308
	0.2	0.2	0.048	0.081	0.551	0.052	0.050	0.061	0.102	0.053	0.308
2000	0	0	0.054	0.068	0.530	0.047	0.048	0.068	0.091	0.062	0.500
	0	0.2	0.047	0.070	0.515	0.051	0.032	0.057	0.090	0.057	0.488
	0.2	0	0.048	0.073	0.505	0.036	0.047	0.054	0.086	0.059	0.512
	0.2	0.2	0.052	0.070	0.473	0.060	0.042	0.072	0.123	0.062	0.478
10000	0	0	0.054	0.087	0.539	0.052	0.052	0.069	0.123	0.062	0.568
	0	0.2	0.055	0.076	0.549	0.052	0.056	0.063	0.094	0.050	0.576
	0.2	0	0.049	0.071	0.533	0.052	0.053	0.066	0.104	0.056	0.561
	0.2	0.2	0.050	0.066	0.535	0.048	0.055	0.069	0.109	0.043	0.561

Notes: 1. MA(1) DGP: $y_t = \mu_y + \beta_y t + u_{yt};\ u_{yt} = \varepsilon_{yt} - \theta_y \varepsilon_{yt-1},$
$x_t = \mu_x + \beta_x t + u_{xt};\ u_{xt} = \varepsilon_{xt} - \theta_x \varepsilon_{xt-1}.$

2. I(d) DGP: $y_t = \mu_y + \beta_y t + u_{yt};\ (1-L)^d u_{yt} = \varepsilon_{yt},$
$x_t = \mu_x + \beta_x t + u_{xt};\ (1-L)^d u_{xt} = \varepsilon_{xt}.$

Table 3. Proportion of rejections ($|t_{\hat{\gamma}}| > 1.96$) based on the regression $y_t = \hat{\alpha} + \hat{\gamma} x_t + e_t$ when y_t and x_t are independent nonstationary processes with drifts.

T	β_y	β_x	R.W	I(d) $d = 0.5$
50	0	0	0.685	0.224
	0	0.2	0.750	0.544
	0.2	0	0.747	0.512
	0.2	0.2	0.817	1.000
100	0	0	0.756	0.325
	0	0.2	0.842	0.629
	0.2	0	0.852	0.608
	0.2	0.2	0.950	1.000
500	0	0	0.902	0.585
	0	0.2	0.936	0.802
	0.2	0	0.940	0.826
	0.2	0.2	1.000	1.000
2000	0	0	0.936	0.738
	0	0.2	0.969	0.884
	0.2	0	0.969	0.878
	0.2	0.2	1.00	1.000
10000	0	0	0.970	0.781
	0	0.2	0.988	0.901
	0.2	0	0.989	0.888
	0.2	0.2	1.000	1.000

Notes: 1. R.W. DGP: $y_t = \mu_y + \beta_y t + u_{yt};\ u_{yt} = u_{yt-1} + \varepsilon_{yt},$
$x_t = \mu_x + \beta_x t + u_{xt};\ u_{xt} = u_{xt-1} + \varepsilon_{xt}.$

2. I(d) DGP: $y_t = \mu_y + \beta_y t + u_{yt};\ (1-L)^d u_{yt} = \varepsilon_{yt},$
$x_t = \mu_x + \beta_x t + u_{xt};\ (1-L)^d u_{xt} = \varepsilon_{xt}.$

Table 4. Proportion of rejections ($|t_{\hat{\gamma}}| > 1.96$) based on the regression $y_t = \hat{\alpha} + \hat{\beta}t + \hat{\gamma}x_t + e_t$ when y_t and x_t are independent nonstationary processes with drifts.

T	β_y	β_x	*R.W.*	I(d) $d = 0.5$
50	0	0	0.526	0.180
	0	0.2	0.497	0.194
	0.2	0	0.522	0.171
	0.2	0.2	0.488	0.158
100	0	0	0.655	0.259
	0	0.2	0.652	0.242
	0.2	0	0.645	0.259
	0.2	0.2	0.642	0.238
500	0	0	0.834	0.490
	0	0.2	0.848	0.454
	0.2	0	0.841	0.485
	0.2	0.2	0.846	0.493
2000	0	0	0.931	0.693
	0	0.2	0.926	0.683
	0.2	0	0.928	0.690
	0.2	0.2	0.913	0.672
10000	0	0	0.975	0.774
	0	0.2	0.964	0.778
	0.2	0	0.975	0.755
	0.2	0.2	0.962	0.767

Notes: 1. R.W. DGP: $y_t = \mu_y + \beta_y t + u_{yt};\ u_{yt} = u_{yt-1} + \varepsilon_{yt},$
$x_t = \mu_x + \beta_x t + u_{xt};\ u_{xt} = u_{xt-1} + \varepsilon_{xt}.$

2. I(d) DGP: $y_t = \mu_y + \beta_y t + u_{yt};\ (1 - L)^d u_{yt} = \varepsilon_{yt},$
$x_t = \mu_x + \beta_x t + u_{xt};\ (1 - L)^d u_{xt} = \varepsilon_{xt}.$