Economics Letters 150 (2017) 142–145

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

On spurious regressions with partial unit root processes

Yundong Tu

Guanghua School of Management and Center for Statistical Science, Peking University, Beijing, 100871, China

HIGHLIGHTS

- This paper considers spurious regressions with partial unit root and near partial unit root processes.
- Three models are investigated via simulations to study the rejection probability of the associated t-statistics in detecting the spurious correlation.
- The results reveal that the spurious regression is a not-so-spurious problem, which cannot be eliminated by adding AR(1) errors in re-estimation.
- The advocated balanced regression approach is found effective in detecting the non-existent relationship.

ARTICLE INFO

Article history:
Received 17 August 2016
Received in revised form 18 November 2016
Accepted 20 November 2016
Available online 25 November 2016

JEL classification:
C12
C32
C52

Keywords:
Balanced regression
Partial near unit root
Spurious regression

ABSTRACT

This paper considers spurious regressions with partial unit root and near partial unit root processes. Three models are investigated via simulations to study the rejection probability of the associated t-statistics in detecting the spurious correlation. The results reveal that the spurious regression is a not-so-spurious problem, which cannot be eliminated by adding AR(1) errors in re-estimation. Moreover, the advocated balanced regression approach is found effective in detecting the non-existent relationship.

1. Introduction

The spurious regression between independent variables has attracted a lot of attention since the seminal work by Granger and Newbold (1974) and Phillips (1986). See Ventosa-Santaulària (2009), Zhang (2013), and Chen, Phillips and Tu (2016) for brief reviews of recent theoretical studies and empirical findings. In particular, Chen, Phillips and Tu (2016) and Lin and Tu (2016) demonstrate that spurious regression occurs in models that involve processes moderately deviated from unity. They propose a balanced regression model in which the induced t-statistic for testing spurious correlation is asymptotically standard normal. As a result, statistical inference for spurious correlation becomes practically easy to implement.

The current paper aims to examine the spurious regression between independent (near) partial unit root processes. The (near) partial unit root process (Caner and Hansen, 2001) is a nonlinear transition autoregression model between two regimes, where one regime is generated by a (near) unit root and the other is generated by a mean-reverting process. Zhang (2013) reports simulation evidences that the spurious regression phenomenon occurs with independent partial unit root processes but it can be eliminated when regressions are re-estimated by adding an AR(1) term in the errors, following the suggestion of McCallum (2010). As concluded in Martínez-Rivera and Ventosa-Santaulària (2012), adding AR(1) term in the errors "should not be conceived of as a universal solution to the spurious inference problem". Instead, we show that the balanced regression approach advocated by Chen, Phillips and Tu (2016), and Lin and Tu (2016), is an attractive alternative. In fact, our simulation studies demonstrate that the findings reported in Zhang (2013) are not robust to model designs.

The rest of the paper is structured as follows. Section 2 presents three models that involve regressions with independent (near) partial unit root processes. Section 3 contains simulation evidence to show that our test based on balanced regression to detect spurious regression has proper size in finite samples. Section 4 concludes and remarks for future research.

E-mail address: yundong.tu@gsm.pku.edu.cn.

http://dx.doi.org/10.1016/j.econlet.2016.11.028
0165-1765/

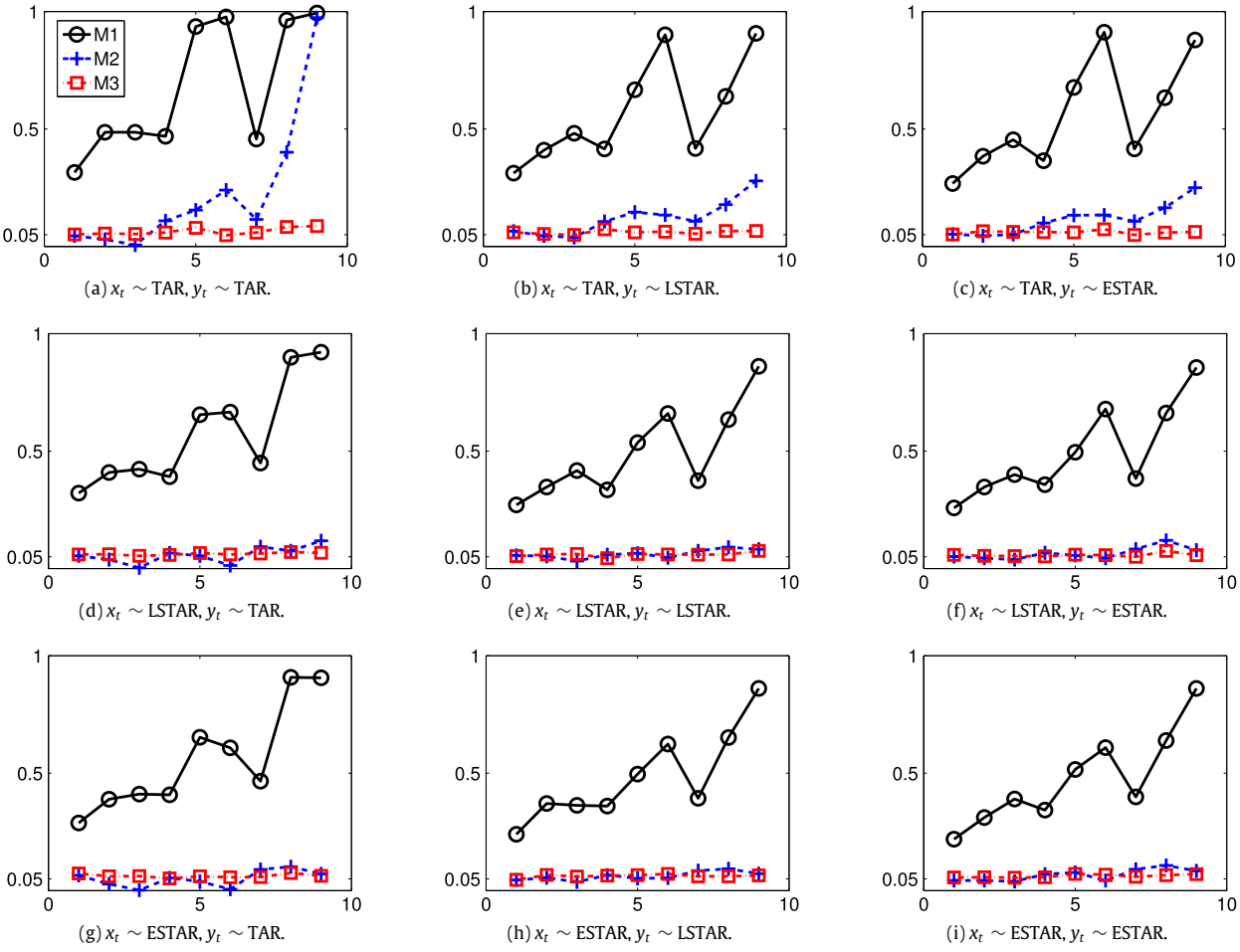

Fig. 1. Rejection frequencies for t-statistics ($|t_{\hat{\beta}_i}| > 1.96$, $i = 1, 2, 3$) in M1, M2 and M3, when x_t and y_t are independent with $n = 50$. The horizontal axes (1–9) in all plots correspond to the nine combinations of c_x and c_y in the order listed in Section 3.

2. The models

We consider the data generating process to capture a transition between two autoregressions:

$$z_t = \mu_z + \theta_z z_{t-1} + \lambda z_{t-1}\pi(z_{t-d}, \kappa, \mu) + u_{zt}, \tag{1}$$

where $\theta_z \geq 0$ and $\lambda < 0$, such that $|\theta + \lambda| < 1$, $\pi(z_{t-d}, \kappa, \mu)$ is a real-valued transition probability function on R^3, in which z_{t-d} is the transition variable, $d \geq 1$, κ and μ are the scale and location parameters, respectively, and u_{zt} is the sequence of zero mean innovations. For the ease of exposition, we shall set $d = 1$ for the subsequent presentation.

We shall consider $\theta_z = 1 + c_z/n$, where c_z is a local parameter and n is the sample size. When $c_z = 0$, we have $\theta_z = 1$ and (1) is called a partial unit root model (Caner and Hansen, 2001; Zhang, 2013), with one extreme regime given by a unit root process and the other a mean-reverting process that represents a linear adjustment toward a long-run equilibrium. When c_z is a constant but nonzero, we call this process a near partial unit root model that includes the nearly nonstationary autoregression ($\lambda = 0$) of Chan and Wei (1987) as a special case.

For the transition function π, three commonly used models are entertained. They include the threshold autoregression (TAR), the logistic smooth transition autoregression (LSTAR), and the exponential smooth transition autoregression (ESTAR), whose transition functions are given below:

$$\textit{TAR}: \quad \pi(z_{t-1}, \mu) = 1\{z_{t-1} \leq \mu\} \tag{2}$$

$$\textit{LSTAR}: \quad \pi(z_{t-1}, \kappa, \mu) = 1/[1 + \exp\{-\kappa(z_{t-1} - \mu)\}] \tag{3}$$

$$\textit{ESTAR}: \quad \pi(z_{t-1}, \kappa, \mu) = 1 - \exp\{-\kappa(z_{t-1} - \mu)^2\}. \tag{4}$$

The spurious regression phenomenon occurs if we regress y_t on x_t when they are generated according to DGP (1) with u_{xt} and u_{yt} being independent. That is, if we consider the OLS regression

$$\text{M1: } y_t = \hat{\alpha}_1 + \hat{\beta}_1 x_t + \hat{\epsilon}_{1t}, \tag{5}$$

the t-statistic associated with $\hat{\beta}_1$ would suggest that y_t and x_t are (spuriously) correlated when standard normal critical values are used.

To tackle this problem, Zhang (2013) follows the idea of McCallum (2010) to consider the following regression instead

$$\text{M2:} \quad \begin{aligned} y_t &= \hat{\alpha}_2 + \hat{\beta}_2 x_t + \hat{\epsilon}_{2t}, \\ \hat{\epsilon}_{2t} &= \hat{\rho}\hat{\epsilon}_{2,t-1} + \hat{e}_t, \end{aligned} \tag{6}$$

where an AR(1) residual is specified. Zhang (2013) illustrates via simulations that the spurious phenomenon can be eliminated in most situations.

Alternatively, we advocate a balanced regression approach to detect spurious regression phenomenon by augmenting (5) with lags x_{t-1} and y_{t-1}, that is,

$$\text{M3: } y_t = \hat{\alpha}_3 + \hat{\beta}_3 x_t + \hat{\gamma}_3 y_{t-1} + \hat{\delta}_3 x_{t-1} + \hat{\epsilon}_3. \tag{7}$$