Spurious Regression and Trending Variables

Antonio E. Noriega and Daniel Ventosa-Santaulària

Escuela de Economía, Universidad de Guanajuato, Dirección General de Investigación Económica, Banco de México

2007

Online at http://mpra.ub.uni-muenchen.de/58775/
MPRA Paper No. 58775, posted 29. September 2014 23:56 UTC
Spurious Regression and Trending Variables

Antonio E. Noriega
Escuela de Economía, Universidad de Guanajuato and
Dirección General de Investigación Económica, Banco de México

Daniel Ventosa-Santaulària
Escuela de Economía, Universidad de Guanajuato

School of Economics Discussion Paper EM07-1
January, 2006

JEL Classification: C12, C13, C22
Keywords: Spurious regression, trends, unit roots, trend stationarity, structural breaks

Abstract

This paper analyses the asymptotic and finite sample implications of different types of nonstationary behavior among the dependent and explanatory variables in a linear spurious regression model. We study cases when the nonstationarity in the dependent and explanatory variables is deterministic as well as stochastic. In particular, we derive the order in probability of the $t$--statistic in a linear regression equation under a variety of empirically relevant data generation processes, and show that the spurious regression phenomenon is present in all cases considered, when at least one of the variables behaves in a nonstationary way. Simulation experiments confirm our asymptotic results.
1 Introduction

It has been documented in recent studies that the phenomenon of spurious regression is present under different forms of nonstationarity in the data generating process (DGP). In particular, when the variables $y_t$ and $x_t$ are nonstationary, independent of each other, ordinary least squares applied to the regression model

$$y_t = \alpha + \delta x_t + u_t$$

have the following implications: 1) the OLS estimator of $\delta$ ($\hat{\delta}$) does not converge to its true value of zero, and 2) the $t$-statistic for testing the null hypothesis $H_0 : \delta = 0$ ($t_\delta$) diverges, thus indicating the presence of an asymptotic spurious relationship between $y_t$ and $x_t$.

The rate at which $t_\delta$ diverges depends on the type of nonstationarity present in the process generating $y_t$ and $x_t$. In Phillips (1986), where a driftless random walk is assumed for both variables, the $t$-statistic is $O_p(T^{1/2})$. For the case of a random walk with drift, Entorf (1997) shows that $t_\delta$ diverges at rate $T$. More recently, Kim, Lee and Newbold (2004) (KLN henceforth) show that the phenomenon of spurious regression is still present even when the nonstationarity in individual series is of a deterministic nature: they find that, under a linear trend stationary assumption for both variables, the $t$-statistic is $O_p(T^{3/2})$. Extending KLN’s results, Noriega and Ventosa-Santaulària (2005) (NVS hereafter), show that adding breaks in the DGP still generates the phenomenon of spurious regression, but at a reduced divergence rate; i.e. $t_\delta$ is $O(T^{1/2})$ under either single or multiple breaks in each variable. In all these works, the implicit assumption is that both variables share the same type of nonstationarity, either stochastic (Phillips, Entorf), or deterministic (KLN, NVS).

Although the literature on this issue has grown considerably, there are still gaps, particularly when regressions involve variables with mixed types of trending mechanisms. The purpose of the present paper is to fill these gaps, using asymptotic and simulation arguments. Our results uncover the presence of spurious regressions under a wide variety of combinations of empirically relevant DGP’s, not explored before in the literature. For instance, we consider regressions of a random walk with drift on a trend (with and without breaks)-stationary process (and vice versa). Section 2 discusses the DGP’s considered. The asymptotic theory developed in Section 3 shows that the rate at which the phenomenon occurs is generally $T^{1/2}$, as predicted by Phillips (1998). However, for some combinations of trending mechanisms the divergence rate is higher. We also show that the spurious regression vanishes when one of the variables is stationary. Section 4 presents some simulation evidence for finite samples, while Section 5 concludes.

---

2 Trending mechanisms in the data generating process

In a simple regression equation, the nature of the trending mechanism in the dependent and explanatory variables is unknown a priori. This is mainly due to a lack of economic knowledge on trending mechanisms. We study the spurious regression phenomenon under eight different DGPs, widely used in applied work in economics.

We consider the following ordinary least squares regression model:

\[ y_t = \alpha + \delta x_t + \epsilon_t \]  

used as a vehicle for testing the null hypothesis \( H_0 : \delta = 0 \). The following assumption summarizes the DGPs considered below for both the dependent and the explanatory variables in model (1).

**Assumption.** The DGPs for \( z_t = y_t, x_t \) are as follows.

<table>
<thead>
<tr>
<th>Case</th>
<th>Name*</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>I(0)</td>
<td>( z_t = \mu_z + u_{zt} )</td>
</tr>
<tr>
<td>2.</td>
<td>I(0)+br</td>
<td>( z_t = \mu_z + \sum_{i=1}^{N_z} \theta_{iz} DU_{izt} + u_{zt} )</td>
</tr>
<tr>
<td>3.</td>
<td>TS</td>
<td>( z_t = \mu_z + \beta_z t + u_{zt} )</td>
</tr>
<tr>
<td>4.</td>
<td>TS+br</td>
<td>( z_t = \mu_z + \sum_{i=1}^{N_z} \theta_{iz} DU_{izt} + \beta_z t + \sum_{i=1}^{M_z} \gamma_{iz} DT_{izt} + u_{zt} )</td>
</tr>
<tr>
<td>5.</td>
<td>I(1)</td>
<td>( \Delta z_t = u_{zt} )</td>
</tr>
<tr>
<td>6.</td>
<td>I(1)+dr</td>
<td>( \Delta z_t = \mu_z + u_{zt} )</td>
</tr>
<tr>
<td>7.</td>
<td>I(1)+dr+br</td>
<td>( \Delta z_t = \mu_z + \sum_{i=1}^{N_z} \theta_{iz} DU_{izt} + u_{zt} )</td>
</tr>
<tr>
<td>8.</td>
<td>I(2)</td>
<td>( \Delta^2 z_t = u_{zt} )</td>
</tr>
</tbody>
</table>

* TS, br, and dr stand for Trend-Stationary, breaks, and drift, respectively.

where \( u_{yt} \) and \( u_{xt} \) are independent innovations obeying the general level conditions of Assumption 1 in Phillips (1986), and \( DU_{izt}, DT_{izt} \) are dummy variables allowing changes in the trend’s level and slope respectively, that is, \( DU_{izt} = 1(t > T_{b_{iz}}) \) and \( DT_{izt} = (t - T_{b_{iz}})1(t > T_{b_{iz}}) \), where \( 1(\cdot) \) is the indicator function, and \( T_{b_{iz}} \) is the unknown date of the \( i^{th} \) break in \( z \). We denote the break fraction as \( \lambda_{iz} = (T_{b_{iz}}/T) \in (0, 1) \), where \( T \) is the sample size.

Note that cases 5, 6 and 7 can be written as

\[
\begin{align*}
z_t & = z_0 + S_{zt} \\
z_t & = z_0 + \mu_z t + S_{zt} \\
z_t & = z_0 + \mu_z t + \sum_{i=1}^{M_z} \theta_{iz} DT_{izt} + S_{zt} \\
\end{align*}
\]

where \( S_{zt} = \sum_{i=1}^{t} u_{zt}, DT_{izt} = \sum_{i=1}^{t} DU_{izt} \), and \( z_0 \) is an initial condition.

Cases 1 and 2 have been used to model the behavior of (theoretically) mean stationary variables, such as real exchange rates, unemployment rates, the great
ratios (i.e. output-capital ratio, consumption-income ratio), and the current account. Examples of \( I(0) \) and \( I(0) \) variables with breaks have been presented in Perron and Vogelsang (1992), Wu (2000), and D’Adda and Scorcu (2003). Cases 3 to 8 are widely used to model growing variables, real and nominal, such as output, consumption, money, prices, etc. Macro variables have been described as \( I(0) \) around a linear trend, \( I(0) \) around a linear trend with structural breaks, and \( I(1) \) in Perron (1992, 1997), Lumsdaine and Papell (1997), Mehl (2000), and Noriega and de Alba (2001). Combinations of case 8 with other cases are often behind the empirical modelling of nominal specifications expressed in terms of \( I(2) \) (nominal) and \( I(1) \) or \( I(0)+ \) breaks (real) variables. Economic models involving \( I(2) \) variables include models of money demand relations, purchasing power parity, and inflation and the markup. Examples of variables described as \( I(2) \) can be found in Juselius (1996, 1999), Haldrup (1998), Muscatelli and Spinelli (2000), Coenen and Vega (2001), and Nielsen (2002).

The DGP's include both deterministic and stochastic trending mechanisms, with 49 possible nonstationary combinations of them among the dependent and the explanatory variables, where case 1 is included as a benchmark.4

The spurious regression phenomenon has already been analyzed for a few combinations of DGP's in the Assumption. For instance, the case of both variables following a unit root (case 5) was studied by Granger and Newbold (1974) and Phillips (1986), and case 6 by Entorf (1997). The case (3) of a trend-stationary model for both variables was studied by Hassler (2000) and KLN, while its extension to multiple breaks (case 4) by NVS. Mixtures of integrated processes were studied by Marmol (1995), who considers cases 5 and 8 (\( y \) follows a unit root, while \( x \) follows two unit roots, and vice versa). Many other combinations, however, have not been analyzed. Among them, combinations 3-6 and 4-6, which have practical importance, given the empirical relevance of structural breaks in the time series properties of many macro variables.

3 Asymptotics for spurious regressions

This section presents the asymptotic behavior of the \( t \)-statistic for testing the null hypothesis \( H_0 : \delta = 0 \) (\( t_\delta \)) in model (1) when the dependent and explanatory variables are generated according to combinations of DGP's in the Assumption.

In the following Theorem, which collects the main results, a combination of DGP's is indicated by the pair \( i - j \), \( (i, j = 1, 2, ..., 8) \) indicating that \( y_t \) is generated by case \( i \), while \( x_t \) by case \( j \), both defined in the Assumption. Thus, for instance, the combination 8 – 5 corresponds to model (1) where \( y_t \) is \( I(2) \) (case 8), while \( x_t \) is \( I(1) \) (case 5).

---

4 We do not consider the cases of \( I(1) \) processes with long memory errors, and fractionally integrated processes, studied in Cappuccio and Lubian (1997), and Marmol (1998), respectively.
Theorem. The order in probability of \( t_\delta \) in model (1) depends on the combination of DGPs for \( y_t \) and \( x_t \) in the Assumption, as follows:

a) \( t_\delta = O_p(1) \) for combination of cases \( 1 - i \) and \( i - 1, i = 1, 2, ..., 8; \)
b) \( t_\delta = O_p(T^{1/2}) \) for combination of cases \( i - i, i = 2, 4, 5, 7, 8, \) and \( i - j; i, j = 2, 3, ..., 8; i \neq j; \) except for combinations \( 3 - 3, 3 - 6, 6 - 3 \) and \( 6 - 6; \)
c) \( t_\delta = O_p(T) \) for combination of cases \( 3 - 6, 6 - 3 \) and \( 6 - 6; \)
d) \( t_\delta = O_p(T^{3/2}) \) for combination of cases \( 3 - 3. \)

Proof. Combination 1-1 represents the classical textbook situation, in which the \( t \)-statistic converges to a standard normal distribution (see for instance White (1984, Chapter V)). Results for combinations 1-2, 2-1, and 2-2 are special cases of Hassler (2003), while 3-3 comes from Hassler (2000) and KLN (who also studied the case 1-3 and 3-1); for 4-4, 5-1, 5-5, 6-6, 8-8, and 5-8 results come from NVS, Hassler (1996), Phillips (1986), Entorf (1997), Marmol (1995), and Marmol (1996), respectively. The proof for the remaining 51 combinations was assisted by the software Mathematica, and, since the mechanics for obtaining the order in probability is the same for each combination of DGPs, we only present the procedure for a single combination, discussed in the Appendix at the end of the paper.

Part a) indicates that when either both variables are \( I(0) \), or one of the variables is \( I(0) \) while the other follows any of the other nonstationary cases, the spurious regression phenomenon is not present, since the \( t \)-statistic does not diverge to infinite; instead, it converges (to a constant, or to a random variable, depending on the DGPs.) For the majority of combination of cases, the \( t \)-statistic diverges (at rate \( \sqrt{T} \) or faster), indicating a spurious relationship among independent variables, as parts b)-d) show.

Table 1 summarizes the above findings. The symmetry of results imply that the order in probability does not depend on the type of nonstationarity among dependent and explanatory variables.

<table>
<thead>
<tr>
<th>DGPs</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>T(^{1/2})</td>
<td>T(^{1/2})</td>
<td>T(^{1/2})</td>
<td>T(^{1/2})</td>
<td>T(^{1/2})</td>
<td>T(^{1/2})</td>
<td>T(^{1/2})</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>T(^{1/2})</td>
<td>T(^{1/2})</td>
<td>T(^{1/2})</td>
<td>T(^{1/2})</td>
<td>T(^{1/2})</td>
<td>T(^{1/2})</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>T(^{1/2})</td>
<td>T(^{1/2})</td>
<td>T(^{1/2})</td>
<td>T(^{1/2})</td>
<td>T(^{1/2})</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>T(^{1/2})</td>
<td>T(^{1/2})</td>
<td>T(^{1/2})</td>
<td>T(^{1/2})</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>T(^{1/2})</td>
<td>T(^{1/2})</td>
<td>T(^{1/2})</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>T(^{1/2})</td>
<td>T(^{1/2})</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>T(^{1/2})</td>
</tr>
</tbody>
</table>
Therefore, when two independent random variables follow any of the nonstationary combinations considered in the Assumption, OLS inference will indicate, asymptotically, a significant (spurious) relationship among them.

The representation theory developed by Phillips (1998) shows that a trending stochastic (or deterministic) process can be represented in various ways. In particular, it can be written as an infinite linear combination of trending deterministic (stochastic) functions with random coefficients. In such an asymptotic environment, he shows that the regression $t$-ratios of the fitted coefficients diverge at rate $\sqrt{T}$. Results from the theorem above indicate that relatively simple nonstationary time series models correctly indicate the presence of the limiting representation.5

4 Experimental results

We computed rejection rates of the $t$-statistic for testing the null hypothesis $H_0 : \delta = 0$, in model (1), using a 1.96 critical value (5% level) for a standard normal distribution. In order to assess the usefulness for a finite sample of the asymptotic results presented in the theorem, rejection rates were based on simulated data, for samples of size $T = 25, 50, 100, 250, 500, 1,000$, and $10,000$, under various combinations of DGPs in the Assumption.6 In all experiments, the number of replications is 10,000.

<table>
<thead>
<tr>
<th>$T$</th>
<th>1-7</th>
<th>2-2</th>
<th>2-4</th>
<th>2-6</th>
<th>3-4</th>
<th>3-6</th>
<th>4-4</th>
<th>4-6</th>
<th>4-7</th>
<th>4-8</th>
<th>6-8</th>
<th>7-8</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>.06</td>
<td>.34</td>
<td>.62</td>
<td>.81</td>
<td>.14</td>
<td>.25</td>
<td>.32</td>
<td>.61</td>
<td>.66</td>
<td>.65</td>
<td>.93</td>
<td>.94</td>
</tr>
<tr>
<td>50</td>
<td>.06</td>
<td>.57</td>
<td>.96</td>
<td>.99</td>
<td>.89</td>
<td>.95</td>
<td>.99</td>
<td>.99</td>
<td>.99</td>
<td>.94</td>
<td>.95</td>
<td>.96</td>
</tr>
<tr>
<td>100</td>
<td>.05</td>
<td>.87</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>.99</td>
<td>1.0</td>
<td>1.0</td>
<td>.97</td>
<td>.97</td>
<td>.97</td>
<td>.97</td>
</tr>
<tr>
<td>250</td>
<td>.05</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>.98</td>
<td>.98</td>
<td>.98</td>
<td>.98</td>
</tr>
<tr>
<td>500</td>
<td>.06</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>1000</td>
<td>.06</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

The values of the parameters in the DGPs are as follows: $\sigma_z = 1$, $\phi_z = 0$, $\mu_x = 0.4$, $\mu_y = 0.7$, $\beta_x = 0.07$, $\beta_y = 0.04$, $\theta_{xz} = 0.07$, $\theta_{yx} = 0.04$, $\gamma_{xi} = 0.02$, $\gamma_{yi} = 0.04$, for $i = 1, \ldots M_z$, $M_z = 2$, for $z = x, y$. Breaks in $x$ ($y$) occur at 20% (40%) and 70% (80%) of total data length.

5The cases of linear deterministic trends without breaks (combinations 3-3, 3-6 and 6-6) do not fit in the general framework of $T^{1/2}$ divergence of Phillips (1998). This is because these cases are not truly spurious, since the limit value of $\delta$ in (1) is not zero, but the ratio of the linear trend parameters in the DGP. On this issue see Hassler (2000) and KLN.

6The experimental results in this section are limited, and only serve as a guide on the finite sample behaviour of some particular cases. The values of the parameters were inspired on real data from Perron and Zhu (1995), comprising historical real per capita GDP series for industrialized economies.
Table 3. Rejection Rates for $t_3$: the case of four breaks

<table>
<thead>
<tr>
<th>Combinations of cases (DGP) in the Assumption</th>
<th>1-7</th>
<th>2-2</th>
<th>2-4</th>
<th>2-6</th>
<th>3-4</th>
<th>3-6</th>
<th>4-4</th>
<th>4-6</th>
<th>4-7</th>
<th>4-8</th>
<th>6-8</th>
<th>7-8</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>.06</td>
<td>.44</td>
<td>.68</td>
<td>.83</td>
<td>.54</td>
<td>.25</td>
<td>.63</td>
<td>.89</td>
<td>.82</td>
<td>.79</td>
<td>.93</td>
<td>.96</td>
</tr>
<tr>
<td>50</td>
<td>.05</td>
<td>.67</td>
<td>.98</td>
<td>.99</td>
<td>.97</td>
<td>.95</td>
<td>.99</td>
<td>.99</td>
<td>.94</td>
<td>.95</td>
<td>.99</td>
<td>.99</td>
</tr>
<tr>
<td>100</td>
<td>.05</td>
<td>.91</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>.97</td>
<td>.97</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>250</td>
<td>.06</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>.98</td>
<td>.98</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>500</td>
<td>.05</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>1000</td>
<td>.05</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Experimental design as above, but with $M_z = 4$, for $z = x, y$. Breaks in $x$ ($y$) occur at 15% (20%), 30% (35%), 50% (55%) and 70% (80%) of total data length.

Tables 2 and 3 present rejection rates under 12 different combinations of DGP's in the Assumption. In Table 2, the cases where breaks are considered (all but 3-6 and 6-8) include 2 breaks, while Table 3 presents results when 4 breaks are allowed. The column labeled 1-7 in both tables presents the finite sample counterpart of the $O_p(1)$ result in part a) of the theorem: the $t$-statistic does not diverge, indicating that the spurious regression phenomenon is not a problem in finite samples. For the rest of cases, the asymptotic nonsense relationship reported in the theorem is also detected in our simulation experiments: the spurious regression phenomenon is present even for samples as small as 25. In comparing results from Tables 2 and 3, it can be noted that a nonsense regression is more likely when the number of structural breaks increases in the DGP.

5 Conclusions

This paper has presented an asymptotic and experimental analysis of the spurious regression phenomenon under a wide variety of empirically relevant data generating processes in a simple regression model. It fills many gaps left open by previous research in the area. In particular, it shows that the $t$-statistic for testing a linear relationship among independent time series diverges, if both variables show a nonstationary behavior, due to either stochastic (unit roots) or deterministic (linear trends and/or structural breaks) factors.

Our results particularize Phillips’ (1998) general results to empirically useful models, by showing that the phenomenon of spurious regression is present for time series with relatively simple trending mechanisms. This phenomenon depends on the commonality of trends and/or breaks in both dependent and explanatory variables. If the nonstationary behavior (stochastic or deterministic) is present in only one of the variables, however, the spurious regression vanishes. Our simulation experiments reveal that a spurious regression is not exclusively an asymptotic phenomenon, it will also be present in finite samples for the majority of DGP's considered.

Given that trending mechanisms as the ones analyzed here are a common feature of the long-run behaviour in many macroeconomic time series, the main conclusion points to a warning regarding inferences drawn from OLS regression.
analysis: the probability of finding a nonsense correlation among independent series will not only be high in finite samples, but it will grow with the sample size.

6 Appendix

We present a guide on how to obtain the order in probability of one combination of DGP s, namely, the combination 1-7, for which

\[ y_t = \mu_y + u_{gt} \]

\[ x_t = x_0 + \mu_x t + \sum_{i=1}^{M_x} \theta_{ix} DT_{ix} + S_{xt} \]

The orders in probability for the rest of cases follow the same steps. The proofs were assisted by the software Mathematica 4.1. The corresponding codes for all combinations of DGP s are available at www.ventosa-santaularia.com/NVS_06a.zip. Below, we describe the steps involved in the computerized calculations.

Write the regression model \( y_t = \alpha + \delta x_t + u_t \) in matrix form: \( y = X \beta + u \). The vector of OLS estimators is \( \hat{\beta} = (X'X)^{-1}X'y \), and the t-statistic of interest \( t_3 = \frac{\hat{\alpha} - \alpha}{\hat{\delta}} \), where \( (X'X)^{-1} \) is the 2nd diagonal element of \( (X'X)^{-1} \) and \( \hat{\alpha} = \frac{1}{T} \sum_{t=1}^{T} \hat{y}_t \), \( \hat{\delta} = \frac{1}{T} \sum_{t=1}^{T} \hat{\delta}_t \). \( t_3 \) is a function of the following objects:

\[ \sum_{t=1}^{T} y_t = \mu_y T + \Sigma_{uy} T^{1/2} \]

\[ \sum_{t=1}^{T} y_t^2 = (\mu_y^2 + \Sigma_{u2y}) T + 2\mu_y \Sigma_{u2y} T^{1/2} \]

\[ \sum_{t=1}^{T} x_t = \frac{1}{2} \left[ \mu_x + \sum_{i=1}^{M_x} \theta_{i} (1 - \lambda_i) \right] T + \Sigma_{sx} T^{3/2} + \frac{1}{2} \left[ \mu_x + \sum_{i=1}^{M_x} \theta_{i} (1 - \lambda_i) \right] T \]

\[ \sum_{t=1}^{T} x_t^2 = \frac{1}{2} \mu_x^2 + \lambda^+ + \frac{1}{2} \mu_x \sum_{i=1}^{M_x} \theta_{i} (1 - \lambda_i)^2 (\lambda_i + 2) \]

\[ \sum_{t=1}^{T} y_t x_t = \left( \mu_y (\mu_x + \sum_{i=1}^{M_x} \theta_{i} (1 - \lambda_i)^2) \right) T + O_p(T^{3/2}) \]

with

\[ \Sigma_{uy} = T^{-1/2} \sum_{t=1}^{T} u_{gt} \]

\[ \Sigma_{u2y} = T^{-1} \sum_{t=1}^{T} u_{gt}^2 \]

\[ \Sigma_{sx} = T^{-3/2} \sum_{t=1}^{T} S_{xt} \]

\[ \Sigma_{sx} = T^{-5/2} \sum_{t=1}^{T} tS_{xt} \]

\[ \Sigma_{sx} = T^{-7/2} \sum_{t=1}^{T} tS_{xt} \]

\[ \lambda^+ = \frac{1}{2} \sum_{i=1}^{M_x} \theta_{i}^2 (1 - \lambda_i)^2 + \sum_{i=1}^{M_x} \sum_{j=1}^{M_x} \theta_{i} \theta_{j} \left[ \frac{3}{2} (1 - \lambda_u(i,j)) \right] + \lambda_d(i,j) (1 - \lambda_u(i,j))^2 \]

\[ \lambda_u(i,j) = \max(\lambda_i, \lambda_j), \quad i, j = 1, 2, ..., M_x \]
\[ \lambda_{l(i,j)} = \min(\lambda_i, \lambda_j) \]
\[ \lambda_{d(i,j)} = \lambda_u(i,j) - \lambda_l(i,j) \]

where (see for instance Phillips (1986)),
\[ \Sigma_{uy} \Rightarrow \sigma_y W_y(1) \]
\[ \Sigma_{u2y} \Rightarrow \sigma_y^2 \]
\[ \Sigma_{ux} \Rightarrow \sigma_x \int_0^1 W_x(r)dr \]
\[ \Sigma_{txx} \Rightarrow \sigma_x \int_0^1 rW_x(r)dr \]
\[ \Sigma_{t1xi} \Rightarrow \sigma_x \sum_{i=1}^{M_x} \theta_i \int_{\lambda_i}^{1} (r - \lambda_i)W_x(r)dr, \]

\( \Rightarrow \) signifies convergence in distribution, and \( W_z(r), z = y, x \) is the standard Wiener process on \( r \in [0, 1] \).

Using these expressions, Mathematica computes the limiting distribution of the parameter vector and the rest of the elements of the output by factoring out the relevant expressions in powers of the sample size. In this way, the orders in probability can be determined, and the limiting expression obtained, by retaining only the asymptotically relevant terms, upon a suitable normalization. From Mathematica’s output it can be deduced that, for the case at hand,
\[ T^{3/2} \delta \left[ \sum_{i=1}^{M_x} \theta_i \int_{\lambda_i}^{1} (r - \lambda_i)W_x(r)dr \right] \]

as reported in the Theorem.

7 References


MATHEMATICA 4.1 (2000), Wolfram Research Inc.


